

SEARCH REQUEST FORM

11-206

Requestor's
Name: _____

Serial
Number: _____

Date: _____

Phone: _____

Art Unit: _____

10607

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 11-17-98
 Searcher: PMB x8-4291
 Terminal time: 12
 Elapsed time: prep 18
 CPU time: _____
 Total time: _____
 Number of Searches: _____
 Number of Databases: 11

Search Site

____ STIC
 ____ CM-1
 ____ Pre-S

Type of Search

0 N.A. Sequence
2 A.A. Sequence
 ____ Structure
 ____ Bibliographic

Vendors

____ IG
 ____ STN
 ____ Dialog
 ____ APS
 ____ Geninfo
 ____ SDC
 ____ DARC/Questel
Mr S Other

WORLD

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 22:12:04 1998; MasPar time 1501.65 Seconds
1522.943 Million cell updates/sec

Tabular output not generated.

Title: >US-08-887-977-9

Description: (1-1119) from US08887977.seq

Perfect Score: 1119

N.A. Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCCCTCAA 1119
Comp: TACAAAGCTGAGTGCACTT.....TCTTTCGACTCAGAGGATT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 552174 seqs, 1021863385 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155

1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro

Database: genbank107
12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pri 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 10.987; Variance 4.759; scale 2.308

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1104	98.7	1255	21	HSCKRL3	H. sapiens G protein-coupled	0.00e+00
2	1104	98.7	3693	21	HSU45984	Homo sapiens CCR6 chem	0.00e+00
3	1100	98.3	1518	21	HSU68032	Human G protein-couple	0.00e+00
4	1100	98.3	2978	21	HSU68030	Human G protein-couple	0.00e+00
5	1070	95.6	1137	20	HSU60000	Human IL8-related rece	0.00e+00
6	530	47.4	1302	22	AB009369	Mus musculus mRNA for	0.00e+00
7	75	6.7	1500	20	HSDNABLR2	H. sapiens BLR2 gene.	1.42e-39
8	75	6.7	2139	20	HUMEBII1CDN	Human G protein-couple	1.42e-39
9	75	6.7	2154	20	HUMGPCRA	Human Epstein-Barr vir	1.42e-39
10	75	6.7	2215	20	HUMEBII103	Human G protein-couple	1.42e-39
11	75	6.7	133801	21	AC004595	Homo sapiens chromosom	1.42e-39
12	64	5.7	2577	20	HSU45982	Human G protein-couple	3.27e-30
13	61	5.5	1495	20	HUMCCCR1A	Human C-C chemokine re	1.04e-27
14	61	5.5	1609	17	E13385	CDNA encoding human MI	1.04e-27
15	62	5.5	2072	22	MUSEBII1CDN	Mouse G protein-couple	1.54e-28

16	5.5	2156	17	158541	Sequence 1 from patent	1.04e-27
17	5.5	2156	20	HUMRANTES	Homo sapiens macrophag	1.04e-27
18	5.5	2214	21	HUMHML45	Human mRNA for HML45.	1.04e-27
19	5.3	528	22	MMLCRI1	Mus musculus mRNA enco	4.72e-26
20	5.3	1180	22	MMLCRI2	Mus musculus lcr-1 gen	4.72e-26
21	5.3	1223	22	MMU59760	Mus musculus fusin hom	4.72e-26
22	5.3	1422	22	MMLCRI3	Mus musculus lcr-1 gen	4.72e-26
23	5.3	1877	22	D87747	Mouse mRNA for murine	4.72e-26
24	5.3	3366	22	MMU65580	Mus musculus fusin (CX	4.72e-26
25	5.2	1050	22	RNU90610	Rattus norvegicus CXC	3.14e-25
26	5.2	7218	17	I66494	Sequence 14 from paten	3.14e-25
27	5.1	1068	15	RAB118C	Oryctolagus cuniculus	2.08e-24
28	5.1	1068	21	AF017282	Macaca mulatta chemoki	2.08e-24
29	5.1	1232	15	RAB118REC	Oryctolagus cuniculus	2.08e-24
30	5.1	1575	22	MMLSTRPT	M. musculus mRNA for le	2.08e-24
31	5.1	1656	16	AB010713	Cyprinus carpio mRNA f	2.08e-24
32	5.1	1744	15	BTU19947	Bos taurus interleukin	2.08e-24
33	5.1	1809	22	AB000803	Mouse mRNA for murine	2.08e-24
34	5.1	3770	22	MMLSTRGN	M. musculus gene encodi	2.08e-24
35	4.9	1068	21	AF026535	Homo sapiens chemokine	8.88e-23
36	4.9	1201	20	HSU28694	Human eosinophil CC ch	8.88e-23
37	4.9	1689	20	HSU49727	Human C-C chemokine re	8.88e-23
38	4.9	1717	21	HSU51241	Human eosinophil eotax	8.88e-23
39	4.7	1068	21	MMV13776	Macaca mulatta CCR-3 g	3.68e-21
40	4.7	1068	21	AF017283	Macaca mulatta chemoki	3.68e-21
41	4.6	1083	17	E13909	CDNA encoding human MC	1.47e-19
42	4.6	1083	21	HUMMCP1R	Human mRNA for monocy	1.47e-19
43	4.6	1140	15	FCU92795	Felis catus fusin (CXC	2.34e-20
44	4.6	1979	20	HSU03905	Human monocyte chemoa	1.47e-19
45	4.6	2232	20	HSU03882	Human monocyte chemoa	1.47e-19

ALIGNMENTS

1	LOCUS	HSCKRL3	1255 bp	DNA	PRI	13-NOV-1996
1	DEFINITION	H.sapiens G protein-coupled receptor	CCR-L3.			
1	ACCESSION	Z79784				
1	NID	g1668737				
1	KEYWORDS	G Protein-coupled Receptor CCR-L3.				
1	SOURCE	human.				
1	ORGANISM	Homo sapiens				
1	REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1255)				
1	AUTHORS	Gutierrez, J., Varona, R., Zaballos, A., Lind, P. and Marquez, G.				
1	TITLE	unpublished				
1	JOURNAL	Unpublished				
1	REFERENCE	2 (bases 1 to 1255)				
1	AUTHORS	Zaballos, A.				
1	TITLE	Direct Submission				
1	JOURNAL	Submitted (03-SEP-1996) Angel Zaballos, Research, Pharmacia & Upjohn, Antonio Lopez 109, Madrid, 28026, Spain				
1	REFERENCE	3 (bases 1 to 1255)				
1	AUTHORS	Zaballos, A., Varona, R., Gutierrez, J., Lind, P. and Marquez, G.				
1	TITLE	Molecular cloning and RNA expression of two new human chemokine receptor-like genes				
1	JOURNAL	Biochem. Biophys. Res. Commun. 227 (3), 846-853 (1996)				
1	MEDLINE	97040707				
1	FEATURES	Location/Qualifiers				
1	source	1..1255				
1		/organism="Homo sapiens"				
1		/db_xref="taxon:9606"				
1		/chromosome="6"				
1		76..1185				
1		/codon_start=1				
1		/product="G PROTEIN-COUPLED RECEPTOR CCR-L3"				
1		/db_xref="PID:e264774"				
1		/db_xref="PID:g1668738"				
1		/translation="MNFSDVDSSEDFVSVNTSYVSDSEMLLCSLQEVRFSLRVF				
1		PLAYSLLICVFLGLNLLVITFAFYKRSMTDYLLNLAIDLEFLVLPFWAVSHA				
1		TGAWFSNATCKLGIYAINFNCGMLLTCTISMDRYIAIVQATKRSRLTPRSK				
1		IICLVVWGLSVIISSTFTFNQKYNQTSQDCEPKYQTVSEPIRWKLLMLGLELFGF				

FIPLMFIFCYTTFIVKTLVQONSKRHKALRVIIIAVVLVFLACQIPHNVMVLTAAHL
GKNRSCOSEKLLGYFTVTEVLAFHCCCLNPVLYAFIGKFRNYFLKILKDLKMCVR
KYKSGFSCAGRISSENISROTSETADNDASSFTM

BASE COUNT 302 a 284 c 284 g 375 t
ORIGIN

Query Match 98.7%; Score 1104; DB 21; Length 1255;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 1114; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 89 ATGTTTTCGACTCCAGTGAAGATATTTTGTGTCAGTCAATACATCTCATATTACTCAGTTG 148
QY 1 ATGTTTTCGACTCCAGTGAAGATATTTTGTGTCAGTCAATACATCTCATATTACTCAGTTG 60
Db 149 ATTCTGAGATGTTACTGTCCTCTTGAGGAGGTCAGGCAGTTCTCCAGGCTATTTGTAC 208
QY 61 ATTCTGAGATGTTACTGTCCTCTTGAGGAGGTCAGGCAGTTCTCCAGGCTATTTGTAC 120
Db 209 CGA-TTGCCCTACTCCTTGATGTCGTTTGTGCGCTCCTGGGGAATATTTCTGGTGGTGATC 267
QY 121 CGAATTGCCCTACTCCTTGATGTCGTTTGTGCGCTCCTGGGGAATATTTCTGGTGGTGATC 180
Db 268 ACCTTTGCTTTTATAGAAAGCCAGGTCCTATGACAGACGTCCTATCTTGAACATGGCC 327
QY 181 ACCTTTGCTTTTATAGAAAGCCAGGTCCTATGACAGACGTCCTATCTTGAACATGGCC 240
Db 328 ATTGCAGACATCCTTTTGTGTTTACTCTCCCATCTCGGCGAGTGAGTCATGCCACTGGT 387
QY 241 ATTGCAGACATCCTTTTGTGTTTACTCTCCCATCTCGGCGAGTGAGTCATGCCACTGGT 300
Db 388 GCCTGGGTTTTCAGCAATGCCAGTCGCAAGTGTCTAAAGCATCTATGCCATCACTTT 447
QY 301 GCCTGGGTTTTCAGCAATGCCAGTCGCAAGTGTCTAAAGCATCTATGCCATCACTTT 360
Db 448 AACTGGGGATGCTGCTCCTGACTTGTCATGACGACGTCATGACGACGTCATGCCATGTACAG 507
QY 361 AACTGGGGATGCTGCTCCTGACTTGTCATGACGACGTCATGACGACGTCATGCCATGTACAG 420
Db 508 GCGACTAAGTCATTCGGGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCCTT 567
QY 421 GCGACTAAGTCATTCGGGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCCTT 480
Db 568 GTTGTGGGGGCTGCAGTCATCATCTCAGCTCAACTTTTGTCTTCAACCAAAATAC 627
QY 481 GTTGTGGGGGCTGCAGTCATCATCTCAGCTCAACTTTTGTCTTCAACCAAAATAC 540
Db 628 AACACCAAGGCGAGCATGTCGTGAACCAAGTACCAGACTGTCCTGGAGCCCATCAGG 687
QY 541 AACACCAAGGCGAGCATGTCGTGAACCAAGTACCAGACTGTCCTGGAGCCCATCAGG 600
Db 688 TGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGTTTCTTTATCCCTTTGATGTTTC 747
QY 601 TGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGTTTCTTTATCCCTTTGATGTTTC 660
Db 748 ATGATATTTTGTACAGTTTCATGTCGCAACCTGTTGCAAGTCAGAAATCTTAAAGG 807
QY 661 ATGATATTTTGTACAGTTTCATGTCGCAACCTGTTGCAAGTCAGAAATCTTAAAGG 720
Db 808 CACAAAGCCATCGTGTATCATAGCTGTGGTGTGTTGTTTCTTCTTCTGCTTGTGCATTCCT 867
QY 721 CACAAAGCCATCGTGTATCATAGCTGTGGTGTGTTTCTTCTTCTGCTTGTGCATTCCT 780
Db 868 CATAACATGCTGCTGTTGAGGGCTGCAAAATTTGGTGAATAAAGAACCGATCTGCCAG 927
QY 781 CATAACATGCTGCTGTTGAGGGCTGCAAAATTTGGTGAATAAAGAACCGATCTGCCAG 840
Db 928 AGCGAAAAGCTAATTTGGCTATACGAAATCTGCAGAAAGTCCTGCTGCTTCTTCTGCTGCTG 987
QY 841 AGCGAAAAGCTAATTTGGCTATACGAAATCTGCAGAAAGTCCTGCTGCTTCTTCTGCTGCTG 900
Db 988 TGCCTGAACCCCTGCTGCTACGCTTTTATTTGGCGAGAGTTCAGAAACTACTTTCTGAAG 1047
QY 901 TGCCTGAACCCCTGCTGCTACGCTTTTATTTGGCGAGAGTTCAGAAACTACTTTCTGAAG 960

Db 1048 ATCTTGAAGGACCTGTGTGTGTGTGAGAGGAAGTACAAAGTCTCTCAGGCTTCTCCTGTGCC 1107
QY 961 ATCTTGAAGGACCTGTGTGTGTGTGAGAGGAAGTACAAAGTCTCTCAGGCTTCTCCTGTGCC 1020
Db 1108 GGGAGGTACTCAGAAAACATTTCTGGCAGACCACTGAGACCGCAGATACGACATGCG 1167
QY 1021 GGGAGGTACTCAGAAAACATTTCTGGCAGACCACTGAGACCGCAGATACGACATGCG 1080
Db 1168 TCGCTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1206
QY 1081 TCGCTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 2

LOCUS HSU45984 3693 bp DNA PRI 09-JUL-1997
DEFINITION Homo sapiens CCR6 chemokine receptor (CMKR6) gene, complete cds.
ACCESSION U45984
NID 92246432
KEYWORDS human.
SOURCE human.

ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 3693)

AUTHORS Baba, M., Imai, T., Nishimura, M., Kakizaki, M., Takagi, S.,
Hieshima, K., Nomiya, H. and Yoshie, O.

TITLE Identification of CCR6, the specific receptor for a novel

lymphocyte-directed CC chemokine LARC

J. Biol. Chem. 272 (23), 14893-14898 (1997)

JOURNAL MEDLINE 97313465

REFERENCE 2 (bases 1 to 3693)

AUTHORS Lautens, L.L., Modi, W. and Bonner, T.I.

TITLE Cloning, Tissue Distribution and Chromosomal Localization of a

Potential G-Protein-Linked Receptor

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 3693)

AUTHORS Bonner, T.I.

TITLE Direct Submission

Submitted (18-JAN-1996) Tom I. Bonner, Lab of Cell Biology, NIMH,

Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA

JOURNAL Location/Qualifiers

FEATURES

source

1. 3693

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="6q27"

/clone="GPR-CY4"

join(35..140,237..3136)

132..1352

/gene="CMKR6"

join(132..140,237..1352)

/gene="CMKR6"

/note="G protein-coupled receptor"

/codon_start=1

/product="CCR6 chemokine receptor"

/db_xref="PID:G2251211"

/translation="MSGESMNFSDVDSSEDFVSVNTSYYSVSEMLLSLQEVRFQ
SRLEFVPLAYSLLICVGLLNVITFAFKKARSMTDYLNNMADILFVLTPW
AVSHATGAWFESNATCKLLKGIYAINFNCGMLLTICISMDRYIAIVQATKFRLSRT
LPRSKIIICLVVWGLSVIISSTFVNQKYSTGGSDVCEPKYQTVSEPRKLLMLGLE
LLFGFFPLMFIMFYCTFIVKTLVQONSKRHKALRVIIIAVVLVFLACQIPHNVMVLT
TAANLGNRSCOSEKLLGYFTVTEVLAFHCCCLNPVLYAFIGKFRNYFLKILKDL
KVCVRKYKSGFSCAGRISSENISROTSETADNDASSFTM"

polyA_signal

polyA_site 3137

BASE COUNT 1015 a 726 c 828 g 1124 t

ORIGIN

Query Match 98.7%; Score 1104; DB 21; Length 3693;

Best Local Similarity 99.6%; Pred. No. 0.00e+00;

Matches 1114; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 1335 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1373
|||||
QY 1081 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 3
LOCUS H5U68032 1518 bp DNA PRI 06-MAR-1997
DEFINITION Human G protein-coupled receptor (STRL22) gene, complete cds.
ACCESSION U68032
NID g1870668
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Liao, F., Lee, H. H. and Farber, J. M.
TITLE Cloning of STRL22, a new human gene encoding a G-protein-coupled receptor related to chemokine receptors and located on chromosome 6q27
JOURNAL Genomics 40 (1), 175-180 (1997)
MEDLINE 9724503
REFERENCE 2 (bases 1 to 1518)
AUTHORS Farber, J. M. and Liao, F.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1996) NIAID/Bldg10/Rm11n-228, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..1518
/organism="Homo sapiens"
/note="Caucasian"
/db_xref="taxon:9606"
/clone_lib="Stratagene # 951202"
/chromosome="6"
/map="6q27"
/sex="male"
1..1518
/gene="STRL22"
intron <1..2
/gene="STRL22"
exon 3..108
/gene="STRL22"
join(100..108,205..1320)
/gene="STRL22"
/note="contains seven transmembrane domains; expressed in lymphocytes"
/codon_start=1
/product="G protein-coupled receptor"
/db_xref="PID:g1870669"
/translation="MSGSMNFDVDFDSSDYFVSNTSYSDSEMLCSLQEVRF
SRLEVPVIAVSLICVGLGNLIVITFAFYKARSMTDVYLLNMAIDILFVLTLPFW
AVSHATGAWFVSNAATCKLLGIYAINFCGMILLTCTISMDRYIAIVQATKSFRLSRT
LPRTKILIVWGLSVIISSSFVFNQKYNQSDVCEPKQIVSEPKWKLMLGLE
LIFGFFPLMFIFCYTFIVKTLVQANSKRHKAIRVIAVLVFLPACQIPNMLLV
TAANGLKMRNSQSEKLYITVTVLAFVLCCLNPLVLYAFVIGQKFRNYFLKILKDL
WCVRRYKSGSFSCAGRYSENISROTSETADNDNASFTM"
109..204
intron
/gene="STRL22"
exon 205..>1518
/gene="STRL22"
BASE COUNT 360 a 354 c 351 g 453 t
ORIGIN

Query Match 98.3%; Score 1100; DB 21; Length 1518;
Best Local Similarity 99.4%; Pred. No. 0.00e+00;
Matches 1112; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Db 224 ATGTTTCGACTCCAGTGAAGATTATTTGTGTGTCAGTCAATCTTATATCTACGTTG 283
|||||
QY 1 ATGTTTCGACTCCAGTGAAGATTATTTGTGTGTCAGTCAATCTTATATCTACGTTG 60
|||||
Db 284 ATCTGAGATGTTACTGTCTTCCTTCGAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTAC 343
|||||

Db 256 ATGTTTCGACTCCAGTGAAGATTATTTGTGTGTCAGTCAATCTTATATCTACGTTG 315
|||||
QY 1 ATGTTTCGACTCCAGTGAAGATTATTTGTGTGTCAGTCAATCTTATATCTACGTTG 60
|||||

Db 316 ATCTGAGATGTTACTGTCTCTCTTCGAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTAC 375
|||||
QY 61 ATCTGAGATGTTACTGTCTCTCTTCGAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTAC 120
|||||

Db 376 CGA-TTGCCTACTCTGATCTGTCTTTTGGCCTCCTGGGGAATATCTGTGGTGTATC 434
|||||
QY 121 CGAATTGCCTACTCTGATCTGTCTTTTGGCCTCCTGGGGAATATCTGTGGTGTATC 180
|||||

Db 435 ACCTTTGTCTTTTAAAGAGCCAGGCTATGACAGAGCTATCTCTTGACATGGCC 494
|||||
QY 181 ACCTTTGTCTTTTAAAGAGCCAGGCTATGACAGAGCTATCTCTTGACATGGCC 240
|||||

Db 495 ATTCGACAGATCTCTTTTGTCTTACTCTCCCATCTTGGGCAGTGAAGTATGCCACTGGT 554
|||||
QY 241 ATTCGACAGATCTCTTTTGTCTTACTCTCCCATCTTGGGCAGTGAAGTATGCCACTGGT 300
|||||

Db 555 GCGTGGGTTTTCAGCAATGCCAGCTGCAAGTTGCTTAAAGGCAATCTATGCCATCAACTTT 614
|||||
QY 301 GCGTGGGTTTTCAGCAATGCCAGCTGCAAGTTGCTTAAAGGCAATCTATGCCATCAACTTT 360
|||||

Db 615 AACTGCGGGATGCTCTCTGATCTGATTCATGATGATGATGATGATGATGATGATGATG 674
|||||
QY 361 AACTGCGGGATGCTCTCTGATCTGATTCATGATGATGATGATGATGATGATGATGATG 420
|||||

Db 675 GCGACTAAGTCAATCCGGCTCCGATCCAGAACATACCCGAGCAAAATCATCTCGCTT 734
|||||
QY 421 GCGACTAAGTCAATCCGGCTCCGATCCAGAACATACCCGAGCAAAATCATCTCGCTT 480
|||||

Db 735 GTTGTGTGGGGCTGTAGTCAATCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 794
|||||
QY 481 GTTGTGTGGGGCTGTAGTCAATCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 540
|||||

Db 795 AACACCCAGGAGGATGCTGTGAACCAAGTACCAGACTGTCTCGGAGCCCATCAGG 854
|||||
QY 541 AACACCCAGGAGGATGCTGTGAACCAAGTACCAGACTGTCTCGGAGCCCATCAGG 600
|||||

Db 855 TGGAGCTGTCTGTGGGGCTGAGTCACTCTTGGTGTCTTGTCTTATCCCTTTGATGTTTC 914
|||||
QY 601 TGGAGCTGTCTGTGGGGCTGAGTCACTCTTGGTGTCTTGTCTTATCCCTTTGATGTTTC 660
|||||

Db 915 ATGATATTTTGTACACGTTTCAATGCTCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG 974
|||||
QY 661 ATGATATTTTGTACACGTTTCAATGCTCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG 720
|||||

Db 975 CACAAAGCCATCCGTTGTAATCATAGCTGTGGTGTCTTGTCTTGTCTGCTGTGATTCCT 1034
|||||
QY 721 CACAAAGCCATCCGTTGTAATCATAGCTGTGGTGTCTTGTCTTGTCTGCTGTGATTCCT 780
|||||

Db 1035 CATACATGGTCTGCTGTGACGGCTCAATTTGGGTTAAATGAAATGAAATGAAATGAAATG 1094
|||||
QY 781 CATACATGGTCTGCTGTGACGGCTCAATTTGGGTTAAATGAAATGAAATGAAATGAAATG 840
|||||

Db 1095 AGCGAAAGCTAATTTGGCTATACGAAACTGTCAAGAGTCTCTGGCTTCTGCTGACTGC 1154
|||||
QY 841 AGCGAAAGCTAATTTGGCTATACGAAACTGTCAAGAGTCTCTGGCTTCTGCTGACTGC 900
|||||

Db 1155 TGCCTGAACCTGTCTACGCTTTTATTTGGGAGAGTTCAGAACTTCTGTAAG 1214
|||||
QY 901 TGCCTGAACCTGTCTACGCTTTTATTTGGGAGAGTTCAGAACTTCTGTAAG 960
|||||

Db 1215 ATCTTGAAGGACCTGTGTGTGTGAGAGGAGTCAAGTCTCTCGGCTTCTCTGTGCC 1274
|||||
QY 961 ATCTTGAAGGACCTGTGTGTGTGAGAGGAGTCAAGTCTCTCGGCTTCTCTGTGCC 1020
|||||

Db 1275 GGGAGGTACTCAGAAACATTTCTGGCAGACCATGAGACCGCAGATTAACGACATGGG 1334
|||||
QY 1021 GGGAGGTACTCAGAAACATTTCTGGCAGACCATGAGACCGCAGATTAACGACATGGG 1080
|||||

Db	592	CGGTGGGTTTTCAGCAATGCCAGTGCAGTGTCTGCTAAAGGTCATCTATGCCATCAACTTT	651			
QY	301	CGGTGGGTTTTCAGCAATGCCAGTGCAGTGTCTGCTAAAGGTCATCTATGCCATCAACTTT	360			
Db	652	AACTCGGGATGCTGCTCCTGACTTGCATATAGCATGGACCGGTACATCGCCATGTACAG	711			
QY	361	AACTCGGGATGCTGCTCCTGACTTGCATATAGCATGGACCGGTACATCGCCATGTACAG	420			
Db	712	CGGACTAAGTCATTCGGGCTCCGATCCAGAACACTACCGCGACGAAATCATCTGCCCTT	771			
QY	421	CGGACTAAGTCATTCGGGCTCCGATCCAGAACACTACCGCGACGAAATCATCTGCCCTT	480			
Db	772	GTGTGTGGGGCTGTCACTCATCTCCAGTCAACTTTTGTCTCAACCAAAAATAC	831			
QY	481	GTGTGTGGGGCTGTCACTCATCTCCAGTCAACTTTTGTCTCAACCAAAAATAC	540			
Db	832	AACACCCCAAGCAGGATGCTGTGAACCCCAAGTACCAAGTGTCTCGGAGCCCATCAGG	891			
QY	541	AACACCCCAAGCAGGATGCTGTGAACCCCAAGTACCAAGTGTCTCGGAGCCCATCAGG	600			
Db	892	TGGAAGTCTGATGTGGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTT	951			
QY	601	TGGAAGTCTGATGTGGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTT	660			
Db	952	ATGATATTTTGTACAGCTTCATTGTCAAAACCTTGGTGCAAGCTCAGAAATCTAAAGG	1011			
QY	661	ATGATATTTTGTACAGCTTCATTGTCAAAACCTTGGTGCAAGCTCAGAAATCTAAAGG	720			
Db	1012	CACAAAGCCATCCGTGTATCATAGCTGTGGTGTCTTCTTCTGCTTGTGTCAGATTCCT	1071			
QY	721	CACAAAGCCATCCGTGTATCATAGCTGTGGTGTCTTCTTCTGCTTGTGTCAGATTCCT	780			
Db	1072	CATAACATGGTCTGCTGTGACGGCTGCAAAATTTGGGTAAATGAACCGATCCCTGCCAG	1131			
QY	781	CATAACATGGTCTGCTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCCCTGCCAG	840			
Db	1132	ACGAAAGACTAATTTGGCTATACGAAACTGTACAGAAAGTCTCTGGCTTTCCTGCACTGC	1191			
QY	841	ACGAAAGACTAATTTGGCTATACGAAACTGTACAGAAAGTCTCTGGCTTTCCTGCACTGC	900			
Db	1192	TGCTTGAAACCTGTCTACGCTTTTATTGGGAGAACTTCAGAAACTACTTTCTGGAAG	1251			
QY	901	TGCTTGAAACCTGTCTACGCTTTTATTGGGAGAACTTCAGAAACTACTTTCTGGAAG	960			
Db	1252	ATCTTGAAAGGACTCTGTGTGTGAGAGGAAGTACAAGTCTCAGGCTTCTCTGTGCC	1311			
QY	961	ATCTTGAAAGGACTCTGTGTGTGAGAGGAAGTACAAGTCTCAGGCTTCTCTGTGCC	1020			
Db	1312	GGGAGGTACTCAGAAACATTTCTCGGAGACCACTGAGACCCAGATACGACAAATGG	1371			
QY	1021	GGGAGGTACTCAGAAACATTTCTCGGAGACCACTGAGACCCAGATACGACAAATGG	1080			
Db	1372	TCGTCTTCACTATGTATAGAAAGTGTAGTCTCCCTAA	1410			
QY	1081	TCGTCTTCACTATGTATAGAAAGTGTAGTCTCCCTAA	1119			
RESULT	5	HSU60000	1137 bp	mrna	PRI	31-AUG-1996
LOCUS						
DEFINITION						Human IL8-related receptor (DRY6) mRNA, complete cds.
ACCESSION						U60000
NID						91515434
KEYWORDS						human.
SOURCE						human.
ORGANISM						Homo sapiens
REFERENCE						Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS						1 (bases 1 to 1137)
TITLE						McCoY, R. and Perlmutter, D.H.
JOURNAL						Cloning of novel IL8-related receptors from hepatic tissue
REFERENCE						Unpublished
AUTHORS						2 (bases 1 to 1137)
						McCoY, R. and Perlmutter, D.H.

TITLE	Direct Submission
JOURNAL	Submitted (28-MAY-1996) Pediatrics, Washington University in St. Louis, 4942 Parkview Place, St Louis, MO 63110, USA
FEATURES	Location/Qualifiers
source	1. 11137
gene	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/cell_line="hepatoma HepG2"
CDS	1. 11137
	/gene="DRY6"
	/gene="DRY6"
	/codon_start=1
	/product="IL8-related receptor"
	/db_xref="PID:91515435"
	/translation="MGSEMSFSDVSEDFVSVNTSYSDSEMLLCSLOEVRF
	SRLEFPIAYSLICLVFALLGNILVITFAFNKARSMDFVFNMAADILFVLTLPFW
	AVSHATGAWFNSATKLLGIVAINFNCMLLTICISMDRYTAIVQATKSFRLRSRT
	LPKSKICLVVWGLSVIISSSVFVNOKYNTLGSDCPEKYQVTSVPRKLLMLGLE
	LLGFPFPLMFIFCYTFIVKILVQAKSRKRAIRVIAVLFVLAQCIPHNVLIV
	TAANLGRNRSQSEKLGITKIVTEVLAFLHCLNPNVLYAFIGKFRNYFLKILKDL
	WCYRRYKSSGFGSCAGRYSENISRTSETADNDNAVVLHYVIES"
BASE COUNT	263 a 270 c 266 g 338 t
ORIGIN	
Query Match	95.6%; Score 1070; DB 20; Length 1137;
Best Local Similarity	98.6%; Pred. No. 0.00e+00;
Matches	1095; Conservative 0; Mismatches 13; Indels 2; Gaps 2;
Db	29 ATGTTTCGACTCCAGTGAAGATTATTGTCAGTCAGTCAATACTTCAATATTACTACCTTG 88
QY	1 ATGTTTCGACTCCAGTGAAGATTATTGTCAGTCAGTCAATACTTCAATATTACTACCTTG 60
Db	89 ATTCGAGATGTTACTGTCTCTTCGAGGAGTGCAGTCTTCCAGGCTATTTCTAC 148
QY	61 ATTCGAGATGTTACTGTCTCTTCGAGGAGTGCAGTCTTCCAGGCTATTTGTAC 120
Db	149 CGA--TTGCTTACTCCTTGATCTGTCTTCCCTCTCGGGAATATTTCTGGTGGTATC 207
QY	121 CGAATTGCTTACTCCTTGATCTGTCTTCTGCTCTGCGGAAATATTTCTGGTGGTATC 180
Db	208 ACCTTTCCTTTATAGAGGCGGAGTCTATGACAGCGTCTATCTCGTGAACATGGCC 267
QY	181 ACCTTTCCTTTATAGAGGCGGAGTCTATGACAGCGTCTATCTCTTCTGAAATGGCC 240
Db	268 ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTTCTGGGAGTGAAGTCACTGGT 327
QY	241 ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTTCTGGGAGTGAAGTCACTGGT 300
Db	328 GCGTGGGTTTTCAGCAATGCCAGTGCAGTGTGCTAAAGGATCTATGCGCATCAACTTT 387
QY	301 GCGTGGGTTTTCAGCAATGCCAGTGCAGTGTGCTAAAGGATCTATGCGCATCAACTTT 360
Db	388 AACTGCGGGATGCTGCTCTGACTTGCATTTAGCATGAGCCGGTACATGCGCATGTACAG 447
QY	361 AACTGCGGGATGCTGCTCTGACTTGCATTTAGCATGAGCCGGTACATGCGCATGTACAG 420
Db	448 GCGATTAAGTCAATCCGCGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 507
QY	421 GCGATTAAGTCAATCCGCGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 480
Db	508 GTTGTGTGGGGCTGTGTCAGTCAATCATCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC 567
QY	481 GTTGTGTGGGGCTGTGTCAGTCAATCATCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC 540
Db	568 AACACCCCTTGGCAGGATGCTGTGAACCCCAAGTACAGACTGTCTCGGTGCCCATCAGG 627
QY	541 AACACCCCAAGGAGGATGCTGTGAACCCCAAGTACAGACTGTCTCGGAGCCCATCAGG 600
Db	628 TGGNAGTCTGTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTTCCCTTTGATGTTT 687
QY	601 TGGNAGTCTGTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTTATCCCTTTGATGTTT 660

Direct Submission
Submitted (28-MAY-1996) Pediatrics, Washington University in St. Louis, 4942 Parkview Place, St Louis, MO 63110, USA

Location/Qualifiers
1. 11137

organism="Homo sapiens"
/db_xref="taxon:9606"

cell_line="hepatoma HepG2"

gene="DRY6"

gene="DRY6"

codon_start=1

product="IL8-related receptor"

db_xref="PID:91515435"

translation="MGSEMSFSDVSEDFVSVNTSYSDSEMLLCSLOEVRF

SRLEFPIAYSLICLVFALLGNILVITFAFNKARSMDFVFNMAADILFVLTLPFW

AVSHATGAWFNSATKLLGIVAINFNCMLLTICISMDRYTAIVQATKSFRLRSRT

LPKSKICLVVWGLSVIISSSVFVNOKYNTLGSDCPEKYQVTSVPRKLLMLGLE

LLGFPFPLMFIFCYTFIVKILVQAKSRKRAIRVIAVLFVLAQCIPHNVLIV

TAANLGRNRSQSEKLGITKIVTEVLAFLHCLNPNVLYAFIGKFRNYFLKILKDL

WCYRRYKSSGFGSCAGRYSENISRTSETADNDNAVVLHYVIES"

BASE COUNT 263 a 270 c 266 g 338 t

ORIGIN

Query Match 95.6%; Score 1070; DB 20; Length 1137;

Best Local Similarity 98.6%; Pred. No. 0.00e+00;

Matches 1095; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

Db 29 ATGTTTCGACTCCAGTGAAGATTATTGTCAGTCAGTCAATACTTCAATATTACTACCTTG 88

QY 1 ATGTTTCGACTCCAGTGAAGATTATTGTCAGTCAGTCAATACTTCAATATTACTACCTTG 60

Db 89 ATTCGAGATGTTACTGTCTCTTCGAGGAGTGCAGTCTTCCAGGCTATTTCTAC 148

QY 61 ATTCGAGATGTTACTGTCTCTTCGAGGAGTGCAGTCTTCCAGGCTATTTGTAC 120

Db 149 CGA--TTGCTTACTCCTTGATCTGTCTTCCCTCTCGGGAATATTTCTGGTGGTATC 207

QY 121 CGAATTGCTTACTCCTTGATCTGTCTTCTGCTCTGCGGAAATATTTCTGGTGGTATC 180

Db 208 ACCTTTCCTTTATAGAGGCGGAGTCTATGACAGCGTCTATCTCGTGAACATGGCC 267

QY 181 ACCTTTCCTTTATAGAGGCGGAGTCTATGACAGCGTCTATCTCTTCTGAAATGGCC 240

Db 268 ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTTCTGGGAGTGAAGTCACTGGT 327

QY 241 ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTTCTGGGAGTGAAGTCACTGGT 300

Db 328 GCGTGGGTTTTCAGCAATGCCAGTGCAGTGTGCTAAAGGATCTATGCGCATCAACTTT 387

QY 301 GCGTGGGTTTTCAGCAATGCCAGTGCAGTGTGCTAAAGGATCTATGCGCATCAACTTT 360

Db 388 AACTGCGGGATGCTGCTCTGACTTGCATTTAGCATGAGCCGGTACATGCGCATGTACAG 447

QY 361 AACTGCGGGATGCTGCTCTGACTTGCATTTAGCATGAGCCGGTACATGCGCATGTACAG 420

Db 448 GCGATTAAGTCAATCCGCGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 507

QY 421 GCGATTAAGTCAATCCGCGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 480

Db 508 GTTGTGTGGGGCTGTGTCAGTCAATCATCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC 567

QY 481 GTTGTGTGGGGCTGTGTCAGTCAATCATCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC 540

Db 568 AACACCCCTTGGCAGGATGCTGTGAACCCCAAGTACAGACTGTCTCGGTGCCCATCAGG 627

QY 541 AACACCCCAAGGAGGATGCTGTGAACCCCAAGTACAGACTGTCTCGGAGCCCATCAGG 600

Db 628 TGGNAGTCTGTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTTCCCTTTGATGTTT 687

QY 601 TGGNAGTCTGTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTTATCCCTTTGATGTTT 660

```

688 ATGATATTTTGTACGTTTCATTTGTCACAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 747
|||||
561 ATGATATTTTGTACGTTTCATTTGTCACAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 720
|||||
748 CACAAAGCCATCCGTGTAATCATAGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCT 807
|||||
721 CACAAAGCCATCCGTGTAATCATAGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCT 780
|||||
808 CATAACATGTCCTGCTGTGTGACGGCTGCAAAATTTGGTGAATGAAACCGATCTGCGCAG 867
|||||
781 CATAACATGTCCTGCTGTGTGACGGCTGCTAATTTGGTGAATGAAACCGATCTGCGCAG 840
|||||
868 AGCGAAAGCTAATTTGCTATACGAACACTGTCAGAAAGTCTGCGCTTCTGCGCTTCTGCGCT 927
|||||
841 AGCGAAAGCTAATTTGCTATACGAACACTGTCAGAAAGTCTGCGCTTCTGCGCTTCTGCGCT 900
|||||
928 TGCTGAACCTGTGCTGTGCTTATTTGGGAGAAAGTCTCAGAAACTTCTTCTGAAG 987
|||||
901 TGCTGAACCTGTGCTGTGCTTATTTGGGAGAAAGTCTCAGAAACTTCTTCTGAAG 960
|||||
988 ATCTTGAAGACCTGTGCTGTGAGAAAGTACAAAGTCTCAGGCTTCTCCTGTGCC 1047
|||||
961 ATCTTGAAGACCTGTGCTGTGAGAAAGTACAAAGTCTCAGGCTTCTCCTGTGCC 1020
|||||
1048 GCGAGTACTCAGAAACATTTCTCGCAGACCACTGAGACCGGAGATACGACAAATGCC 1107
|||||
1021 GCGAGTACTCAGAAACATTTCTCGCAGACCACTGAGACCGGAGATACGACAAATGCC - 1079
|||||
1108 GTGCTCTTCACTATGTGATAGAAGCTGA 1137
|||||
1080 GTGCTCTTCACTATGTGATAGAAGCTGA 1109
|||||

RESULT 6
LOCUS AB009369 1302 bp mRNA ROD 09-DEC-1997
DEFINITION Mus musculus mRNA for G protein-coupled receptor KY411, complete cds
ACCESSION AB009369
NID 92668409
KEYWORDS G protein-coupled receptor KY411.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1302)
Yanagihara.S., Komura.E. and Yamaguchi.Y.
Mouse G protein-coupled receptor KY411.
Published Only in DataBase (1997) in press
2 (bases 1 to 1302)
Yanagihara.S., Komura.E. and Yamaguchi.Y.
Direct Submission
Submitted (01-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Shigehiro Yanagihara, Kirin Brewery Co., Ltd., Pharmaceutical
Research Laboratory, 3 Miyahara-cho, Takasaki, Gunma 370-12, Japan
(E-mail:syanagihara@kirin.co.jp, tel:81-273-46-9826)
Location/Qualifiers
1. .1302
/organism="Mus musculus"
/db_xref="taxon:10090"
147. .1250
/codon_start=1
/product="G protein-coupled receptor KY411"
/db_xref="PID:d1024681"
/db_xref="PID:g2668410"
/translation="MNSTSYFGTDDYDNTYYISIPDHGPCSILEEVRNFKVFPVPIA
YSLICVFLGLNIMVWTFAPFKARSDVDVLLNMAITDILFVLPFWATHTATNT
WYFDALCKLMGTIYAVNFCNGLLACISMDRYIAIVQATKSFVRVSRLLTHSKVIC
VAWVFSIIISPTFFNKKYEQDVCPEPRYSRVSPTITWKLGLGSLFFGFPTP
LLFMVFCYLIIFKTLVQAQNSKRRAIRVIAVLVFLAQIIPHNVLVLTAVNTGV
GRSCTEKVLATRNVAEVLAFHCLCNPVLYAFIGQKPRNYFMKIMKDWCMRRKNK
MPGFLCARYSESYISQRTSEVFNENASFTM"
BASE COUNT 314 a 331 c 320 g 337 t
ORIGIN
```

```

Query Match 47.4%; Score 530; DB 22; Length 1302;
Best Local Similarity 76.9%; Pred. No. 0.00e+00;
Matches 786; Conservative 0; Mismatches 232; Indels 4; Gaps 3;

Db 228 TGCTCCCTAGAGAGGTGACAGAACTTCACCAAGGPTATTTGTGC-AATTGCTTACTCCTT 286
|||||
QY 78 TGCTCCTTCGAGAGGTGACGAGTTCCTCAGGCTATTTGTACCGAATGCTTACTTACTCCTT 137
|||||
Db 287 AATATGTGCTTTTGGCTCCTCTGGCAACATTTATGGTGGTATGACCTTTGCCCTTCTTACAA 346
|||||
QY 138 GATCTGTGTTTGGCTCCTCTGGGAATATTTCTGGTGGTATGACCTTTGCTTTTATAA 197
|||||
Db 347 GAAAGCCAGATCCATGACGTGACGTCTTACCTGTTGAACATGGCCATCAGACATATCTT 406
|||||
QY 198 GAAGCCAGGTCTATGACAGAGCTCTATCTTTGAACATGGCCATTCGACACATCTCTT 257
|||||
Db 407 TGCTCTCAACCTTACCGTCTCTGGCAGTTACTCATGCCACCAACACTTGGTGTTCAGCGA 466
|||||
QY 258 TGTTCTTACTCTCCCATCTCTGGCAGTGAGTCATGCCACTGGTGGCTGCTTTTTCAGCAA 317
|||||
Db 467 TGCACCTGTGTAACACTGATGAAAGGACATATGCGGTCAACTTTAACTGTGGGATGCTGCT 526
|||||
QY 318 TGCCAGTGCAGTTGCTTAAAGGCACTCTATGCCATCACTTTAACTGCGGATGCTGCT 377
|||||
Db 527 CTTGCTGTATCAGATGGACCGGTACATTCGTCATCTGTCAGGCAACAAATCTTTTCOG 586
|||||
QY 378 CTTGCTGTATCAGATGGACCGGTACATTCGTCATCTGTCAGGCAACAAATCTTTTCOG 437
|||||
Db 587 GGTAGCTCCAGAACACTGACGACAGTAAGTCACTCTGTGGCAGTGGTGTTCATCTC 646
|||||
QY 438 GCTCCGATCCAGAACACTTACCAGGCAAGAAATCATCTGCTTGTGTGGGGGTGTC 497
|||||
Db 647 CATCATCTCTCAAGCCCTACATTTATCTTCAACAGAAATACGAGCTGCGAGTATCGTGA 706
|||||
QY 498 AGTCATCTCTCCAGTCAACTTTTGTCTTCAACCAAAATACACACCAAGGACGCA 557
|||||
Db 707 TGTCTGTAGCCACCGGTACAGTCTGTCTCAGAGCCCATCAGTGAAGTGTGCTGGGTAT 766
|||||
QY 558 TGTCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGAAGTGTGATGTT 617
|||||
Db 767 GGGACTGGAGCTGTCTTTGGGTCTTACCCTTGTCTGCTTGTGCTTATGCTTCTGCTATCT 826
|||||
QY 618 GGGGCTTGAGCTTACTCTTTGGTCTTTTATCCCTTTGATGTTGTCATGATTTTGTACAC 677
|||||
Db 827 GTTCATTTATCAGACTTGGTGCAGGCCAGAACTCCAAGAGCCACAGAGCCATCCGAGT 886
|||||
QY 678 GTTCATTTATCAGACTTGGTGCAGACTCAGAAATTTAAAGGCCAAGCCATCCGCTGT 737
|||||
Db 887 CTTGATCGCTGTGTTCTGCTTCTGCTTGTTCAGATCCCTTCAACATGCTGCTCTCT 946
|||||
QY 738 AATCATAGCTGTGCTGCTTGTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 797
|||||
Db 947 CTTGACTGGGTTCACACGGGCAAGTGGGCGGAGCTGACGACCGAGAAAGTCTCTCGC 1006
|||||
QY 798 TGTGAGGCTGCTAATTTGGGTAAATGAACCGATCTCTGCCAGAGCGAAAGTAATGG 857
|||||
Db 1007 CTACACAGGACGTTGGCGGAGTCTGCTGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 1066
|||||
QY 858 CTATACGAAACTGTCACAGAACTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 917
|||||
Db 1067 GTATCGGTTTATGGACAGAAATTCAGAACTACTTTCATGAAGATCATGAAGATGTGTG 1126
|||||
QY 918 CTACGCTTTTATGGCAGAACTTCAAGAACTACTTTCATGAAGATCTTGAAGGACCTGTG 977
|||||
Db 1127 GTGATAGAGAGAGATAAAGATGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186
|||||
QY 978 GTGTGTGAGAGAAAGTACAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
|||||
Db 1187 CTACATCTCCAGGACAGCCAGTGAGACCTCGAAATATGATGATGATGATGATGATGATGATG 1246
|||||
QY 1038 C-ATTTC- -GGCAGACCGAGTGGAGACCGCAGATACGACAAATGCGCTGCTTCTACTAT 1094
|||||
```



```
QY 256 TTTGTTCTTACTCTCCCATCTTGGCAGTGAAGTCATGCCACTGGTGGGTTTTCAGC 315
Db 442 GTCCACTTTTGCAGCTCACTTTTGGCATCTACAGATGAGCTTCTTCAGTGGCATGTC 501
QY 316 AATGCCACGTGCAAGTGTGAAAAGGCATCTATCCATCACTTTAACTGGCGGATGCTG 375
Db 502 CTACTCTTTTGCATCAGCATGACCGTACGTGGCCATCGTCCAGGC 548
QY 376 CTCTGACTTGCATTAGCATGGACCGGTATACGCCATTCATACAGGC 422

RESULT 9 HUMGPCRA 2154 bp mRNA PRI 31-DEC-1994
LOCUS Human Epstein-Barr virus induced G-protein coupled receptor mRNA,
DEFINITION complete cds.
ACCESSION L08176
NID 9183484
KEYWORDS Epstein-Barr virus induced gene; G-protein coupled receptor.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
REFERENCE 1 (bases 1 to 2154)
AUTHORS Birkenbach,M., Josefsen,K., Valamanchili,R., Lencir,G. and Kieff,E.
TITLE Epstein-Barr virus-induced genes: first lymphocyte-specific G
protein-coupled peptide receptors
J. Virol. 67 (4), 2209-2220 (1993)
93188173
FEATURES
Location/Qualifiers
1..2154
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="BL41/B95-8"
/cell_type="B lymphocyte, EBV-converted Burkitt lymphoma"
/germline
/map="17q12-21.2"
64..135
sig_peptide
/gene="EBI 1"
/product="EBV induced G-protein coupled receptor"
64..1200
gene
/gene="EBI 1"
64..1200
CDS
/gene="EBI 1"
/note="putative"
/codon_start=1
/product="EBV induced G-protein coupled receptor"
/db_xref="PID:g183485"
/translation="MDLGRPKSVLVVALLVIFQVCLCODEVTDYIGDNTVDYTLF
ESLSKDKVRNFKAWFLPIMYSICFVGLLNGLVLTIIYFRKLTMTDTYLLNLAV
ADILFLTLFPWAYSAKSWVGFHFKLIFAIYKMSFFSGMLLLCISIDRYVAIQ
AVSAHRARVLLISKLSCVSAIATVLSIPELLYSDLQSSSEQAMRCSLITEHYE
AFITIQVAMVIGFLVPLLAMSFYLVIRTLQARNFERNKAIKVIAVAVVVFVFO
LPYNGVLAOTVANENITSSCELSKOLNAYDVITYSLACVRCVNPFLYAFIGVKFR
NDIFKFLDGLCLSQEQLRQSSCRHRRSSMSVEAEITTFSP"
BASE COUNT 486 a 647 c 543 g 478 t
ORIGIN
Query Match 6.7%; Score 75; DB 20; Length 2154;
Best Local Similarity 63.7%; Pred. No. 1.42e-39;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;
Db 203 TGTGCTCCCAAGAGGACGTGCGGAATT-TAAAGCCTGGTTCCTCCATCATGACTTCC 261
QY 76 TGTGCTCCTTGCAGAGGTGACGAGTCTCCAGGCTATTGTACCGAATTCCTACTCC 135
Db 262 ATCATTTGTTTCGTGGCCCTACTGGCAATGGCTGGTCTGTGTGACCTATATCTTTC 321
QY 136 TTGATCTGTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCATCCTTTGCTTTTAT 195
Db 322 AAGAGGCTCAAGACCATGACCGATACCTACTGCTCAACCTGGGGTGGCGAGACATCTC 381
QY 196 AAGAGGCCAGGTCTATGACAGACGCTATCTCTTGACATGGCCATTGCAGACATCTC 255
```

```
Db 382 TTCCTCCTGACCTCCCTTCTCTGGGCTACAGCGGGCAA--GTCC-TGGGTCCTCGGT 438
QY 256 TTTGTTCTTACTCTCCCAFTTGGGCAGTGAAGTCATGCCACTGGTGGGTTTTCAGC 315
Db 439 GTCCACTTTTGAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCAGTGGCATGTC 498
QY 316 AATGCCACGTGCAAGTGTGAAAAGGCATCTATCCATCACTTTAACTGGCGGATGCTG 375
Db 499 CTACTTCTTTGCATCAGCATGACCGTACGCGCTACGTGGCCATCGTCCAGGC 545
QY 376 CTCTGACTTGCATTAGCATGGACCGGTATACGCCATTCATACAGGC 422

RESULT 10 HUMEBI103 2215 bp DNA PRI 10-AUG-1995
LOCUS Human G protein-coupled receptor (EBI 1) gene exon 3, complete cds.
DEFINITION
ACCESSION L31584 L31539
NID 9468314
KEYWORDS G protein-coupled receptor.
SEGMENT 3 of 3
SOURCE Homo sapiens (tissue library: Statagene #946205) placenta DNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
REFERENCE 1 (bases 1 to 2215)
AUTHORS Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Schweickart,V.L., Raport,C.J., Godiska,R., Byers,M.G., Eddy,R.L.
Jr., Shows,T.B. and Gray,P.W.
TITLE Cloning of human and mouse EBI1, a lymphoid-specific
G-protein-coupled receptor encoded on human chromosome 17q12-q21.2
JOURNAL Genomics 23 (3), 643-650 (1994)
MEDLINE 95154835
FEATURES
Location/Qualifiers
1..2215
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
/tissue_lib="Stratagene #946205"
/map="17q12-21.2"
join(L31582:1376..2586,L31583:1..305,1..2215)
gene
/gene="EBI 1"
join(L31582:1439..1448,L31583:221..270,169..1245)
CDS
/product="G protein-coupled receptor"
/db_xref="PID:g468316"
/translation="MDLGRPKSVLVVALLVIFQVCLCODEVTDYIGDNTVDYTLF
ESLSKDKVRNFKAWFLPIMYSICFVGLLNGLVLTIIYFRKLTMTDTYLLNLAV
ADILFLTLFPWAYSAKSWVGFHFKLIFAIYKMSFFSGMLLLCISIDRYVAIQ
AVSAHRARVLLISKLSCVSAIATVLSIPELLYSDLQSSSEQAMRCSLITEHYE
AFITIQVAMVIGFLVPLLAMSFYLVIRTLQARNFERNKAIKVIAVAVVVFVFO
LPYNGVLAOTVANENITSSCELSKOLNAYDVITYSLACVRCVNPFLYAFIGVKFR
NDIFKFLDGLCLSQEQLRQSSCRHRRSSMSVEAEITTFSP"
introns
/number=2
exons
/number=169..2215
/gene="EBI 1"
BASE COUNT 520 a 665 c 545 g 485 t
ORIGIN
Query Match 6.7%; Score 75; DB 20; Length 2215;
Best Local Similarity 63.7%; Pred. No. 1.42e-39;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;
Db 248 TGTGCTCCCAAGAGGACGTGCGGAATT-TAAAGCCTGGTTCCTCCATCATGACTTCC 306
QY 76 TGTGCTCCTTGCAGAGGTGACGAGTCTCCAGGCTATTGTACCGAATTCCTACTCC 135
Db 307 ATCATTTGTTTCGTGGCCCTACTGGCAATGGCTGGTCTGTGTGACCTATATCTTTC 366
QY 136 TTGATCTGTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCATCCTTTGCTTTTAT 195
Db 367 AAGAGGCTCAAGACCATGACCGATACCTACTGCTCAACCTGGCGGTGGCAGACATCTC 426
```


...
Note: remainder of annotations omitted.

Query Match

```
Best Local Similarity 58.2%; Pred. No. 3.27e-30;
Matches 335; Conservative 0; Mismatches 235; Indels 6; Gaps 6;

Db 265 ATGACCGACATCTCTTTTGAATTTGGCAATTCGTGACCTCTCTTCTTGTCACATCTT 324
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 ATGACAGAGCTATCTCTTGAACTATGGCAATTCGACATTCCTCTTGTCTTACTCTC 270
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 CCGTCTGGCCATT-G-C-TGCTGCTGACCAAGTTCCAGACCTTCATGTGCAAG 381
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CCATCTGGGAGTAGTATGACCATCTGGTGGTGGTTCAGCAATGCACGTGCAAG 330
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GTGGTCAACAGCATGATGAACTTCTACAGCTGTGTGTGCTGATCATGTGCATC 441
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 TTGCTAAAGGCACTATGCACTCAACTTAACTGCGGGATGCTGCTCTGACTTGCAAT 390
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AGCGTGGACAGTATGCACTGCGGAGCCCATGAGACACATCTTGGAGGAGAAA 501
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 391 AGCATGGGCGGTATGCGCCATTTGACAGGCACTAAAGTCAATTCGCGCTCCGATCCAGA 450
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 AGGCTTTTGTACAGCAAAATGTTCTTACCATCTGGTATTGGCAGCTGCTCTGCG 561
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 451 ACACATACCGCGGAGCAAAATCACTGCTGTTGTGTGGGGGCTGTCAGTCACTCATCC 510
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 ATCCAGAAATCTTATACAGCAAAATCAAGAGGAATCCGGCATTTGCTATCTGCACCAATG 621
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 511 AGCTCAACTTTGCTTCAACCAAAATACACACCAAGGCGAGCATGCTGTGACCC 570
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 GTTTACCCCTAGCGA-TGAGAGCACCAAAAC-TGAAGTCAGCTG-TCTTGACCCGAAAGTC 678
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 571 AAGTACCAAACTGTCTCGGAGCCATCAGGTGGAAGTGTGATGTTGGGGCTTGAGCTA 630
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 ATTCTGGGGTCTCTCTCCCTCGTGGTATGCTGCTGCTATACCATCATTCATCAC 738
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 CTCCTTGGTTCTTTATCCCTTGTATGTTGTCATGATATTTTGTACAGCTTCATTTGTCAAA 690
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 ACCCTGATACAAAGCAAGTCTTCCAGACCAAAAGCCCTAAAGTGACCATCAGTC 798
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 ACCTTGGTCAAGTCAGATTCCTAAAGGCAAAAGCCATCCGTTGATCATAGCTGTG 750
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 CTGACCGCTTTTGTCTGTCTCAGTTTCCCTACAAAC 834
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 GTGCTGTGTTCTGCTGCTGATTCCTCATTAAC 786
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
LOCUS HUMCCCKR1A 1495 bp mRNA PRI 31-DEC-1994
DEFINITION Human C-C chemokine receptor type 1 (C-C CKR-1) mRNA, complete cds.
ACCESSION L09230
NID 9179984
KEYWORDS C-C chemokine receptor type 1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1495)
AUTHORS Neote,K., DiGregorio,D., Mak,J.Y., Horuk,R. and Schall,T.J.
TITLE Molecular cloning, functional expression, and signaling
JOURNAL characteristics of a C-C chemokine receptor
MEDLINE Cell 72 (3), 415-425 (1993)
FEATURES
    source
        1..1495
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="differentiated HL60"
    gene
        1..1068
        /gene="CMKR-1"
    CDS
        1..1068
        /gene="CMKR-1"
        /codon_start=1
        /product="C-C chemokine receptor type 1"
        /db_xref="pid:9179985"
        /translation="METPTNTEDYDTTTFDYDATPCQKVNRAFGAQLPLPLSLV

FVIGLVNILVVLVOYKRLKNNMTSIVLLNLAISDLLEFLTFIDYKLVKDDWVFG
DAMKILSGEYVYTGLEYSEIFILLITDRYLAIYHVPALPARYTFFGVTISIIWAL
AILASMPGLRFLSKQWETHHTCSLHPHESLRWKLFQALKNLGLVPLPLLMILIC
ITGLIKILLRPNKSKKAVLFIEMIFELFWPYNLILISVQDFLFTHECQSQS
RHLDAVQVTEVIAYTHCCSNPVIYAFVGERFRYLRQLFHRVAVHLVLYLPFLSV
RLERSVSTSPSTGEHLSAGF"
BASE COUNT 348 a 389 c 361 g 397 t
ORIGIN
Query Match 5.5%; Score 61; DB 20; Length 1495;
Best Local Similarity 60.4%; Pred. No. 1.04e-27;
Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Db 121 TACTCCTTGGTATTGTTCATTTGGCTGTTGGAAACATCCTGGTGGTCTGCTTGTG 180
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 TACTCCTTGGTATTGTTCATTTGGCTGTTGGGAATATCTGGTGGTATCACCTTTGCT 189
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CAATACAAGAGGCTAAATAACATGACAGCATCTACCTCCCTGAACCTGCCATTTCTGAC 240
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 TTTATAGAAGGCCAGGTCTATGACAGACGCTCTATCTTGAACATGCCATGACAGAC 249
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CTGCTCTCTCTTACAGCTTCCCTTCTGATCGACTACAAGTTGAAGGATGACTGGGTT 300
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 ATCTCTTGTCTTACTCTCCCATTTCTGGCAGTGAGTCATGCCCATGTTGCGTGGTT 309
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTTGGTGAAGCATGTGTAAGATCTCTCTGGGTTTATTACACAGGCTTGTACAGCGAG 360
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 TTCAGCATGCCAGTGCAGTTGCTAAAGGCACTATGTCATCACTTAACTGCGGG 369
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ATCTTTTTCATCTCTGCTGACGATTGACAGGTAACCTGGCCATGCTCCAGCC 413
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 ATGCTCTCTGACTTGCATTAGCAGCGGTACATGCCCATGTTACAGGC 422
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
LOCUS E13385 1609 bp DNA PAT 27-APR-1998
DEFINITION cDNA encoding human MIP-1 alpha /RANTES receptor.
ACCESSION E13385
NID 93252190
KEYWORDS JP 1997176048-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1609)
AUTHORS Honda,S. and Fujisawa,T.
TITLE PRODUCTION OF HUMAN MIP-1ALPHA/PANTES RECEPTOR PROTEIN AND USE
JOURNAL Patent: JP 1997176048-A 1 08-JUL-1997;
COMMENT TAKEDA CHEM IND LTD
OS Homo sapiens (human)
PN JP 1997176048-A/1
PD 08-JUL-1997
PF 28-DEC-1995 JP 1995342130
PI HONDA SUSUMU, FUJISAWA TOMOYUKI
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC A61K45/00,
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC A61K45/00,
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC C07K14/705,
PC C12N5/10,C12N15/09,C12P21/02,C12Q1/00,(C12P21/02,C12R1/91), PC
(C12Q1/00,
PC C12R1/91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
PH Key Location/Qualifiers
FT source 1..1609
FT /organism="Homo sapiens"
FT /cell_type="leukemia"
FT /cell_line="U937"
```



```

FT      CDS      /clone='pCCR'
FT      115..1182
FT      /product='human MIP-1 alpha /RANTES receptor'
FT      1..114
FT      1183..1609.
FT      Location/Qualifiers
FEATURES
    source      1..1609
                /organism='Homo sapiens'
                /db_xref='taxon:9606'
BASE COUNT    386 a 416 c 385 g 422 t
ORIGIN
    Query Match      5.5%; Score 61; DB 17; Length 1609;
    Best Local Similarity 60.4%; Pred. No. 1.04e-27;
    Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
    Db 235 TACTCCTTGATTTGTCATTTGGCTGGTGGAAACATCCCTGGTGGTCTGCTGCTTGTG 294
    QY 130 TACTCCTTGATCTGTCTTGTGGCTCTCTGGGGAATATCTTGGTGGTATCACTTTGCT 189
    Db 295 CAATACAAGAGGCTAAACATACACAGCATCTACCTCTGAACTGACCTGGCCATTTCTGAC 354
    QY 190 TTTTATAAGAAGCCAGGCTGTATGACAGACGCTCTATCTTTGAACATGGCCATTCGACAC 249
    Db 355 CTGCTCTCTGTCAGCTTCCCTTGGATCGACTACAAGTTGAAGGATGACTGGGTT 414
    QY 250 ATCCTCTTGTCTTACTCTCCCATCTGGGCACTGATGCTGCTGCTGGTGGT 309
    Db 415 TTTGGTGATGCCATGTGAAGATCTCTCTGGGTTTATTACACAGGCTTGTACAGCGAG 474
    QY 310 TTCAGCAATGCCAGCTGCAAGTCTCTAAAGGCACTATGCGCATCACTTTAACTGCGGG 369
    Db 475 ATCTTTTCAATCCTGCTGCTGACGATTGACAGGTACCTGGCCATCGTCCAGC 527
    QY 370 ATGCTGCTCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAGGC 422

RESULT 15
LOCUS      MUSEB11CDN 2072 bp mRNA ROD 10-AUG-1995
DEFINITION Mouse G protein-coupled receptor (EBI 1) mRNA, complete cds.
ACCESSION  L31580 L31501
NID        g468340
KEYWORDS   G protein-coupled receptor.
SOURCE     Mus musculus (strain B6/CBAFlJ) (library: Stratagene #935303)
           female 6-8 weeks thymus cDNA to mRNA.
ORGANISM   Mus musculus
           Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
           Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
           Murinae; Mus.
REFERENCE  1 (bases 1 to 2072)
AUTHORS   Schweickart,V.L., Raport,C.J., Godiska,R., Byers,M.G., Eddy,R.L.
           Jr., Shows,T.B. and Gray,P.W.
TITLE     Cloning of human and mouse EB1, a lymphoid-specific
           G-protein-coupled receptor encoded on human chromosome 17q12-q21.2
JOURNAL   Genomics 23 (3), 643-650 (1994)
MEDLINE   95154835
FEATURES
    source      Location/Qualifiers
    1..2072
                /organism="Mus musculus"
                /strain="B6/CBAFlJ"
                /db_xref="taxon:10090"
                /dev_stage="6-8 weeks"
                /sex="female"
                /tissue_type="thymus"
                /tissue_lib="Stratagene #935303"
    1..175
                /gene="EBI 1"
    1..2072
                /gene="EBI 1"
    176..1312
                /gene="EBI 1"
    /codon_start=1
    /product="G protein-coupled receptor"

5'UTR
gene
CDS
```

```

/db_xref="PID:g468341"
/translation="MDPKRKNLVWLVIFQVCFQCODEVDDYIGENTVDTLY
ESVCFKDRNFKAFWFLPMYFVGLLGNGLVILTYIFKRLKMTDTYLLNAV
ADILFLLPFWAYSEAKSWIFGYLCRGIFGYKLSFFSGMLLCLISIDRYVAIVQ
AVSRHRHARVLLISKLVGIMLALFPELLYSGLOKNSGDTLRCSVSVAQVE
ALITIOVAQMVFGFLVPLAMSFYLIIRTLQARNFERNKAIKVIILAVVVFVFO
LPYNGVLAQTVANFNINSSCETSQKLNIAIDVTYSLASVRCVNPFLYAFIGVKFR
SDLFKFLKDLGLCSQERLHRHSSCRHRNARSVSMEAEITTTTFSP"
3'UTR
polyA_site
/gene="EBI 1"
2072
BASE COUNT    482 a 585 c 528 g 477 t
ORIGIN
    Query Match      5.5%; Score 62; DB 22; Length 2072;
    Best Local Similarity 63.1%; Pred. No. 1.54e-28;
    Matches 185; Conservative 0; Mismatches 105; Indels 3; Gaps 2;
    Db 368 TATTCTGTCATCTGCTTGGGCGCTGCTCGGCAACGGGCTGGTGATATGAGGTACATC 427
    QY 130 TACTCCTTGATCTGTCTTGGCCTCTCTGGGAATATCTTGGTGGTATCACTTTGCT 189
    Db 428 TATTTCAAGAGGCTCAAGACCATGACGGGATACCTGCTCAACCTGGCCGTGGCAGAC 487
    QY 190 TTTTATAAGAAGCCAGGCTATGACAGACGCTCTATCTTGAACATGGCCATTGCAGAC 249
    Db 488 ATCCTTTTCTCTTAATCTTCCCTTCTGGGCTACAGCGAAGCCAA--GTCC-TGGATC 544
    QY 250 ATCCTCTTTGTTTCTTACTCTCCCATCTCTGGCAGTGAGTCATGCCACTGGTGGTGGT 309
    Db 545 TTTGGCGTCTACTCTGTGAAGGCGATCTTTTGGCATCTATAAGTTAAAGTCTTCTACGCGG 604
    QY 310 TTCAGCAATGCCAGCTGCAAGTTGCTTAAAGGCATCTATGCCATCACTTTAACTGCGGG 369
    Db 605 ATGCTGCTGCTCTCTATGCAATCAGCATTTACCGCTAGGTAGCCATCGTCAGGC 657
    QY 370 ATGCTGCTCTGACTTGCATTAGCATGGACCGGTACATCGCCATTGTACAGGC 422

Search completed: Fri Nov 13 23:15:55 1998
Job time : 3831 secs.
```

WQ5RELH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sat Nov 14 00:17:57 1998; MasPar time 167.79 Seconds
Tabular output not generated. 907.323 Million cell updates/sec

Title: >US-08-887-977-9
Description: (1-1119) from US08887977.seq
Perfect Score: 1119
N.A. Sequence: 1 ARGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119
Comp: TACAAAGCTGAGTGACTT.....TCTTTCGACTCAGAGGATT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.061; Variance 4.959; scale 1.827

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	1118	99.9	1119	40	Human dendritic cell	0.00e+00
2	75	6.7	1900	11	Partial coding sequen	3.74e-34
3	75	6.7	2058	11	Putative seven trans	3.74e-34
4	75	6.7	2154	11	Epstein Barr virus in	3.74e-34
5	75	6.7	2180	11	Putative seven trans	3.74e-34
6	61	5.5	1085	33	Human MIP-1alpha/RAN	1.99e-24
7	61	5.5	1495	11	C-C chemokine recepto	1.99e-24
8	61	5.5	2156	34	Human MIP-1 alpha/RAN	1.99e-24
9	62	5.5	2751	11	Seven transmembrane r	4.11e-25
10	57	5.1	1200	17	Recombinant high affi	1.05e-21
11	57	5.1	1200	5	Sequence encoding a h	1.05e-21
12	55	4.9	1071	36	Human CCR3 chemokine	2.33e-20
13	55	4.9	1116	23	CC-chemokine receptor	2.33e-20

14	55	4.9	1193	23	T31335	CC-chemokine receptor	2.33e-20
15	55	4.9	1689	32	T58783	Human C-C chemokine r	2.33e-20
16	55	4.9	1689	23	T31334	CC-chemokine receptor	2.33e-20
17	55	4.9	1915	34	T85162	Human chemokine recep	2.33e-20
18	55	4.9	5099	38	T93601	Human eosinophil eota	2.33e-20
19	55	4.6	1083	35	T96976	Human monocyte chemoa	1.09e-17
20	51	4.6	1979	15	O96298	Human monocyte chemoa	1.09e-17
21	51	4.6	2232	15	O96297	Human monocyte chemoa	1.09e-17
22	49	4.4	1176	17	Q99950	Recombinant high affi	2.28e-16
23	49	4.4	1373	17	Q99951	Recombinant high affi	2.28e-16
24	49	4.4	1607	23	T35277	Chemokine receptor K5	2.28e-16
25	49	4.4	1883	17	O99006	Interleukin 8 recepto	2.28e-16
26	49	4.4	1933	13	O80520	Interleukin-8 recepto	2.28e-16
27	49	4.4	1933	5	Q29505	Interleukin-8 recepto	2.28e-16
28	48	4.3	1162	11	Q66170	Seven transmembrane r	1.03e-15
29	48	4.3	2254	11	Q66167	Seven transmembrane r	1.03e-15
30	46	4.1	1373	5	Q30013	Sequence encoding a l	2.08e-14
31	45	4.0	1106	5	Q30012	Sequence encoding a l	2.08e-14
32	45	4.0	1510	6	Q38747	Human IL-8 receptor c	9.22e-14
33	45	4.0	1748	17	Q99008	Interleukin 8 recepto	9.22e-14
34	44	3.9	1586	26	T44099	Human G-protein recep	4.06e-13
35	43	3.8	1106	17	Q99952	Recombinant high affi	1.77e-12
36	43	3.8	1317	11	Q66179	Seven transmembrane r	1.77e-12
37	43	3.8	1737	17	Q99007	Chemokine superfamily	1.77e-12
38	43	3.8	1737	5	Q29506	New platelet factor 4	1.77e-12
39	43	3.8	1737	13	Q80521	Human monocyte PF4AR	1.77e-12
40	41	3.7	204	1	N81164	Base substituted E.co	3.31e-11
41	40	3.6	1244	7	Q45656	Murine somatostatin r	1.41e-10
42	40	3.6	1299	11	Q63919	Interleukin 8 (IL-8)	1.41e-10
43	39	3.5	1059	34	T85163	Macaque chemokine rec	5.95e-10
44	39	3.5	1080	21	T30385	Probe based on angiot	5.95e-10
45	38	3.4	514	23	T35278	Chemokine receptor cd	2.48e-09

ALIGNMENTS

RESULT 1
ID V15418 standard; cDNA; 1119 BP.
AC V15418;
DT 11-JUN-1998 (first entry)
DE Human dendritic cell chemokine receptor encoding cDNA.
KW Human; thymus expressed chemokine; TECK; MIP-3alpha; MIP-3beta;
OS Homo sapiens.
FH Key
FT CDS
FT Location/Qualifiers
FT 1..1098
FT /*tag= a
FT /product= "DC CR"
FT /note= "dendritic cell chemokine receptor"
FT 577..579
FT /*tag= b
FT /note= "encodes His or Gln, but is stated as Gln in the protein (shown in W48086)."
FT unsure
PN W09801557-A2.
PD 15-JAN-1998.
PR 02-JUL-1997; U10819.
PR 04-JUN-1997; US-048593.
PR 05-JUL-1996; US-675814.
PR 11-OCT-1996; US-028329.
PA (SCHE) SCHERING CORP.
PI Gish KC, Schall TJ, Vicari A, Wang W, Zlotnik A;
DR WPI; 98-101054/09.
DR P-PSDB: W48086.
PT Novel chemokines, e.g. thymus expressed chemokine - used for treating inflammatory conditions including asthma.
PS Claim 5; Page 92-94; 202pp; English.
CC The present sequence encodes a human dendritic cell chemokine receptor.
CC Antibodies which bind to the protein can be used in detecting or diagnosing various immunological conditions related to expression of the protein. The nucleic acid can be used for screening and CC isolating DNA clones for the chemokines, especially from other CC species. The chemokine can be used in the treatment of conditions CC associated with abnormal physiology or development, including

CC inflammatory conditions such as asthma.
SQ Sequence 1119 BP; 261 A; 268 C; 257 G; 332 T;

Query Match 99.9%; Score 1118; DB 40; Length 1119;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1118; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgttttcgactcagtgagaattatttttggctcagtcgaatacttcattactcagttg 60
Qy 1 ATGTTTTCGACTCCAGTGAAGATTATTTTGTCTAGTCAATATCTCATATTTACTCAGTTG 60

Db 61 attctgagatttactgtctccctgcaggaggtcagcagcttctccaggtctattttgac 120
Qy 61 ATTCGAGATGTTACTGTGCTCCCTGCGAGAGGTGAGGAGTCTCCAGGCTATTGTGTAC 120

Db 121 cgaattgctactccttgatctgtttcttggccctccctggggaatatctgtgtgtgac 180
Qy 121 CGAATTGCCACTCTTGATCTGTGTTCTTGCCCTCCCTGGGGAATATCTGTGTGTGATC 180

Db 181 accttgccttttataagaagccaggtctctatgacagacgtctctcttcttgaaatggcc 240
Qy 181 ACCTTTGCTTTTATAGAGAGCCAGGTCTATGACAGACGCTATCTCTGTAACATGGCC 240

Db 241 attgcagacatccttttcttactctccattctctggcagtgagtcagtcactggt 300
Qy 241 ATTCGACATCTCTTTGTTCTTACTCTCCATTCTGGGAGTGTGATGATGCCACTGGT 300

Db 301 gctgggttttcagcaatgcccagtgcaagttgtctaaaggtcatctatgcatcaacttt 360
Qy 301 GCGTGGGTTTCACCAATGCCACGTGCAAGTGTCTAAAGGCATCTATGCCATCACTTT 360

Db 361 aactgctggatgctgctactgactgactgactgactgactgactgactgactgactgact 420
Qy 361 AACTGCGGGATGCTGCTCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACT 420

Db 421 gcaactaagtcattccggctccgactccagaacactacacgcagcagcaaatcctgctt 480
Qy 421 GCGACTAAGTCTATTCGGCTCCGATCCAGAACACTTACCGCAGCAAAATCATCTGCTT 480

Db 481 gttgtgtggggctgctcagtcactatctccagctcaactttgtctcaacaaataac 540
Qy 481 GTTGTGTGGGGCTGTGACTCATCTCTCCAGCTCACTTTGCTTCAACCAAAATAC 540

Db 541 aacacccaagcagcagatgtgtgtaacccaagtcaccanaactgtctcgagcccatcagg 600
Qy 541 AACACCCAAGCAGCAGATGTCTGTGAACCCAAGTACCAAACTGTCTCGAGGCCATCAGG 600

Db 601 tggagctgctgattgtgggcttgagctactctttgtttctttatcccttgatgttc 660
Qy 601 TGGAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGTTGTTTATCCCTTTGATGTT 660

Db 661 atgatatgttttacagcttcattgtcaaaccttggtgcaagctcagaattctctaaagg 720
Qy 661 ATGATATTTGTTACAGCTTCATTGTCTAAACCTTGGTGCAGCTCAGAACTCTAAAAGG 720

Db 721 cacaagccatccgtgtaataatagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Qy 721 CACAAAGCCATCCGTGTAATCATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780

Db 781 cataacatgctgt 840
Qy 781 CATAAACATGCTGCTGT 840

Db 841 agcgaaagctaattgtgtctatcagaaactgtcagaaagtcctggtcttctcagctgc 900
Qy 841 AGCGAAAGCTAATTGGCTATACGAATACTGTACAGAAGTCCCTGGCTTCTCGACATGC 900

Db 901 tgcctgaacctgtgctctacgctttttattggcgagaagttcagaataactttctgaag 960
Qy 901 TGCTGAACCTGTGCTCTACGCTTTTATTGGCGAGAAGTTCAGAAACTACTTTCTGAAG 960

Db 961 atcttgaagacactgt 1020
Qy 961 ATCTTGAAGACCTGTGTTGT 1020

Qy 961 ATCTTGAAGACCTGTGTTGT 1020

Db 1021 gggaggtactcagaaacattttctcggcagaccagtgagaccgagataacgacaatgcy 1080
Qy 1021 GGGAGGTACTCAGAAACATTTCTCGCAGACCAGTGAGACCAGACGATACGACAAATGCG 1080

Db 1081 tgcgtcctcactatgtgatagaaagctgagctcctcctaa 1119
Qy 1081 TCGTCTTCTACTATGTGATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 2
ID Q66162 standard; DNA; 1900 BP.
AC Q66162;
DE 02-FEB-1995 (first entry)
DT Partial coding sequence of seven transmembrane receptor (V31).
KW primer: seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT intron 1..168
FT exon /*tag- a
FT exon 169..1245
FT exon /*tag- b
FT 3'utr 1246..1900
FT /*tag- d
PN W09412635-A.
PD 09-JUN-1994.
PF 17-NOV-1993: U11153.
PR 17-NOV-1992: US-977452.
PA (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
DR WPI; 94-200264/24.
PT DNA encoding seven transmembrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
PS Example 3; Page 54-56; 100pp; English.
CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA
CC purified from leukocytes. Approximately 1000 clones were isolated
CC after the initial amplification reaction and probed with sequences
CC specific for seven transmembrane receptors IL8R1, AT2R and R20.
CC Clones which did not hybridise were then chosen for sequence
CC analysis. Three new clones were identified that appeared to encode
CC seven transmembrane receptor segments. Two more primers (Q66151,
CC Q66152) were used to isolate a full length version of one of these
CC clones designated V31 (See Q66153). This is the sequence of exon 3
CC of the V31 genomic clone (along with partial intron sequences) and
CC is the final exon of the clone. The TAG stop codon is found at
CC positions 1243-1245 of the sequence.
SQ Sequence 1900 BP; 431 A; 594 C; 458 G; 417 T;

Query Match 6.7%; Score 75; DB 11; Length 1900;
Best Local Similarity 63.7%; Pred. No. 3.74e-34;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 248 tgtgtctccaaagagcagtcggaactt-taaagctggttctcctcctcatgactcc 306
Qy 76 TGTGCTCTCTTGCGAGAGGTGAGGAGTCTTCCAGGCTATTGTGACGAATGTGCTACTCC 135

Db 307 atcaattgtttgt 366
Qy 136 TTGATCTGTGTTCTTGGCCCTCTTGGGGAATATTCTGTGTGTGTGTGTGTGTGTGTGT 195

Db 367 aagagctcaagacacatgacgatacctcactcctcactcactcactcactcactcactc 426
Qy 196 AAGAAGCCAGGCTATGACAGACGCTATCTTGAACATGGCCATTTGACAGATCCTC 255

Db 427 ttctctctgaccttccctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 483
Qy 256 TTGTGTTCTACTCTCCATCTCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 315

Db 484 gtccacttttgaagctcatctttgtgcatctacaaagatgagcttctcagtgagctgc 543
Qy 484 gtccacttttgaagctcatctttgtgcatctacaaagatgagcttctcagtgagctgc 543


```

QY 130 TACTCCTTGAATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189
Db 181 caatacaagaggctaaataacatgacagcagctaccctctgaacctggtcatttttgac 240
QY 190 TTTTATAAGAGGCGGAGGTCTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249
Db 241 ctgctctctctgttcacgcttccctctctggtatcgactacaagtgaagatgactgggtt 300
QY 250 ATCCTCTTTGTTCTTACTCTCCCATCTGCGGAGTGAGTCATGCCACTGGTGGTGGGTT 309
Db 301 ttgtgtgacgcatgttaagatcctctctggtgtttattacacagggctgttacagcgag 360
QY 310 TTCAGCAATGCCAGTCGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369
Db 361 atcttttcatcctcgtcgtcagcagattgacaggtgacccgtggtcgtccacgc 413
QY 370 ATGCTGCTCTGACTTGATTCATGATGACGACCGGTACATGCCATGTACAGGC 422

```

RESULT 8

```

ID T90384 standard; DNA; 2156 BP.
AC T90384;
DT 21-JAN-1998 (first entry)
DE Human MIP-1 alpha/RANTES receptor gene.
KW Macrophage inflammatory protein-1 alpha; MIP-1 alpha;
KW reduced upon activation normal T expressed and secreted; RANTES;
KW receptor; cytokine; antiinflammatory; inflammation; human; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 63..1130 /*tag= a
FT FT
PN US5652133-A.
PD 29-JUL-1997.
PF 28-JAN-1993; 012988.
PR 28-JAN-1993; US-012988.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Murphy PM;
DR WPI; 97-392945/36.
DR P-PSDB; W26388.
PT MIP-1-alpha and RANTES receptor nucleic acid - used to develop
PT products for the detection of these cytokine(s) and their receptors,
PT particularly in inflammatory processes
PS Claim 1; Column 11-16; 12pp; English.
CC This nucleic acid sequence encodes a claimed receptor for macrophage
CC inflammatory protein-1 alpha (MIP-1 alpha) and regulated upon
CC activation normal T expressed and secreted (RANTES) protein. Also
CC claimed are: (1) a subsequence of the nucleic acid, having at least
CC 12 contiguous nucleotides; (2) a cell transformed or transfected
CC with the nucleic acid; and (3) purified MIP-1 alpha/RANTES receptor
CC polypeptide. The products can be used for detecting the MIP-1
CC alpha/RANTES receptor and polymorphisms in physiological samples.
CC In addition, the receptor can be expressed and used to assay for
CC MIP-1a/RANTES in biological samples. The quantitation of MIP-1
CC alpha/RANTES is useful for monitoring the levels of these cytokines
CC in a patient. Such measurements are useful in following the anti-
CC inflammatory effects of drugs and prospective usefulness of new anti-
CC inflammatory agents.
SQ Sequence 2156 BP; 538 A; 547 C; 504 G; 567 T;

```

```

Query Match 5.5%; Score 61; DB 34; Length 2156;
Best Local Similarity 60.4%; Pred. No. 1.99e-24;
Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```

```

Db 183 tactccttggtattgttcattgctggtgttgaaacatcctggtgctggtcctgttggtg 242
QY 130 TACTCCTTGAATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189
Db 243 caatacaagaggctaaataacatgacagcagctaccctctgaacctggtcatttttgac 302
QY 190 TTTTATAAGAGGCGGAGGTCTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249
Db 303 ctgctctctctgttcacgcttccctctctggtatcgactacaagtgaagatgactgggtt 362

```

```

QY 250 ATCCTCTTTGTTCTTACTCTCCCATCTGCGGAGTGAGTCATGCCACTGGTGGGTT 309
Db 363 ttgtgtgacgcatgttaagatcctctctggtgtttattacacaggtgtgttacagcgag 422
QY 310 TTCAGCAATGCCAGTCGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369
Db 423 atcttttcatcctcgtcgtcagcagattgacaggtgacccgtggtcgtccacgc 475
QY 370 ATGCTGCTCTGACTTGATTCATGATGACGCGGTACATGCCATGTACAGGC 422

```

RESULT 9

```

ID Q66164 standard; DNA; 2751 BP.
AC Q66164;
DT 03-FEB-1995 (first entry)
DE Seven transmembrane receptor (V31) coding sequence.
KW Primer; seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT Intron 1..691 /*tag= a
FT Exon 692..1771 /*tag= b
FT poly_a_signal 2341..2348 /*tag= c
FT FT
PN W09412635-A.
PD 09-JUN-1994.
PF 17-NOV-1993; U11153.
PR 17-NOV-1992; US-977452.
PA (ICOS-) ICOS CORP.
DR Godiska R, Gray PW, Schweickart VL;
DR WPI; 94-200264/24.
DR P-PSDB; R53747.
PT DNA encoding seven transmembrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
PS Example 5; Page 58-60; 100pp; English.
CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA
CC purified from leukocytes. Approximately 1000 clones were isolated
CC after the initial amplification reaction and probed with sequences
CC specific for seven transmembrane receptors IL8R1, Ar2R and R20.
CC Clones which did not hybridise were then chosen for sequence
CC analysis. Three new clones were identified that appeared to encode
CC seven transmembrane receptor segments. Two more primers (Q66151,
CC Q66152) were used to isolate a full length version of one of the
CC human V31 clone (See Q66153). A fragment of the human clone was used
CC to isolate this V31 genomic clone of the mouse from a library.
SQ Sequence 2751 BP; 651 A; 747 C; 684 G; 669 T;

```

```

Query Match 5.5%; Score 62; DB 11; Length 2751;
Best Local Similarity 63.1%; Pred. No. 4.11e-25;
Matches 185; Conservative 0; Mismatches 105; Indels 3; Gaps 2;

```

```

Db 824 tattctgtcatctgcttctggtggtcgtcgtcgcaacgggtggtgatactgacgtacatc 883
QY 130 TACTCCTTGAATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189
Db 884 tattcaagaggctcaagaccatgacggatcacctacactgctcaacctggcgtggcagac 943
QY 190 TTTTATAAGAGGCGGAGGTCTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249
Db 944 atcttttctcctaattcttccctctctggtggtcctacagcagaagccaa--gtcc-tggatc 1000
QY 250 ATCCTCTTTGTTCTTACTCTCCCATCTGGGAGTGAGTCATGCCACTGGTGGGTT 309
Db 1001 ttgggtgctcaccctgtgtaaggcgcatcttggcatctataagttaagctcttcacggg 1060
QY 310 TTCAGCAATGCCAGTCGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369
Db 1061 atgctgctgctcctatgcacatcagcattgaccgctagcattgacccatgctccaggc 1113

```

QY 370 ATGCTGCTCTGACTTGCTAGCATGAGCCGGTACATGCCATTGTACAGGC 422

RESULT 10
ID Q99949 standard; DNA; 1200 BP.
AC Q99949;
DT 24-APR-1996 (first entry)
DE Recombinant high affinity interleukin-8 receptor subtype A encoding DNA.
KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
KW adult respiratory distress syndrome; neutrophil detection; ss.
OS Oryctolagus cuniculus.

Key Location/Qualifiers
FT 108..1175
FT cds /*tag= a
FT /product= IL-8A_receptor

PN W09525126-A1.
PD 21-SEP-1995.
PF 09-MAR-1995; U03032.
PR 15-MAR-1994; US-210250.
PR 02-MAY-1994; US-237937.
PA (REPK) REPLIGEN CORP.
PA (UYBO-) UNIV BOSTON.
PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
PI Witt DT;
DR WPI; 95-36945/43.
DR P-PSDB; R80950.
PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
PT useful for treating inflammatory disorders, for detecting
PT neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
PS Claim 2; Fig 1A-B; 74pp; English.
CC Monoclonal antibodies were raised against recombinant interleukin-8
CC (IL-8) receptor subtypes A and B from both human and rabbit sources
CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is
CC a high affinity receptor and the B subtype receptor (IL-8B) is a
CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
CC the IL-8 binding domain thus blocking its activation. The mAbs are
CC useful for treating inflammatory disorders (see key words) and for
CC detecting the presence of neutrophils in a biological sample. The
CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
SQ Sequence 1200 BP; 234 A; 366 C; 295 G; 303 T;

Query Match 5.1%; Score 57; DB 17; Length 1200;
Best Local Similarity 62.1%; Pred. No. 1.05e-21;
Matches 185; Conservative 0; Mismatches 110; Indels 3; Gaps 3;

Db 257 ctatgcccctgctctctgctgagcctgctgggcaactccctggtgatgctgtgctact 316
QY 129 CFATCCTTGATCTGTGTTCTTGGCCCTCTGGGGAATATCTCTGGTGATCACCTTTGC 188
Db 317 gtacagccggagcaacgcttcggttcaccgactctacctgctgaacctggccatggccga 376
QY 189 TTTTATAGAGAGCCAGGCTCTATGACAGAGCTATCTCTTGAACATGCCATTGCAGA 248
Db 377 cctgctcttttgcctgaccatgcctctatctgggcccgtctc-caaggaaaaag-gc-tggat 433
QY 249 CATCCTCTTTGTTCTTACTCTCCCACTTCTGGGCGAGTGAGTCATGCCACTGGTGGGTGGT 308
Db 434 ttccggcacgcccctgtgcaagggtgctcgtctgtggaaggaagcaactctacagtg 493
QY 309 TTTTACGACATGCCAGCTGCAAGTGTCTAAAGGCACTCTATGCCATCAACTTTAACTCGG 368
Db 494 aatcctgctctggcctgacagtgatgacccgactacccctggccattgtccatctact 551
QY 369 GATGCTGCTCTGACTTGCTAGCATGAGCCGGTACATGCCATTGTACAGCGGACT 426

RESULT 11
ID Q30011 standard; cDNA; 1200 BP.
AC Q30011;
DT 04-APR-1993 (first entry)
DE Sequence encoding a high affinity recombinant rabbit interleukin-8

DE (IL-8) receptor polypeptide in F3R.
KW IL-8 receptor polypeptide; G-protein-coupled receptor; ss.
OS Oryctolagus cuniculus.

Key Location/Qualifiers
FT 108..1172
FT cds /*tag= a

PN W09218641-A.
PD 29-OCT-1992.
PF 10-APR-1992; U02977.
PR 10-APR-1991; US-685101.
PR 09-JUL-1991; US-726606.
PR 09-DEC-1991; US-803842.
PA (REPK) REPLIGEN CORP.
PA (UYBO-) UNIV BOSTON.
PI Navarro J, Thomas KM, Witt DP;
PI WPI; 92-382123/46.
DR P-PSDB; R28272.
PT Recombinant mammalian interleukin-8 receptor - used for screening
PT interleukin-8 binding antagonists, used to treat inflammation
PS Disclosure; Fig 1; 71pp; English.
CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to
CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
CC plaques were screened for those which hybridized to an antisense
CC oligonucleotide (Q30015). This probe was designed based on the
CC sequence derived from the second transmembrane domain of G-protein-
CC coupled receptors. After tertiary screening, six plaques were
CC isolated. The insert of one of these plaques, termed F3R was of 2.5
CC kb in size. This insert was sequenced. The protein deduced from
CC the F3R clone demonstrates that it belongs to the family of
CC G-protein-coupled receptors. The deduced protein sequence
CC indicates seven putative transmembrane segments.
SQ Sequence 1200 BP; 234 A; 366 C; 294 G; 304 T;

Query Match 5.1%; Score 57; DB 5; Length 1200;
Best Local Similarity 62.1%; Pred. No. 1.05e-21;
Matches 185; Conservative 0; Mismatches 110; Indels 3; Gaps 3;

Db 257 ctatgcccctgctctctgctgagcctgctgggcaacccctggtgatgctgtgctact 316
QY 129 CTATCCTTGATCTGTGTTCTTGGCCCTCTGGGGAATATCTCTGGTGATCACCTTTGC 188
Db 317 gtacagccggagcaacgcttcggttcaccgactctacctgctgaacctggccatggccga 376
QY 189 TTTTATAGAGAGCCAGGCTCTATGACAGAGCTCTATCTCTTGAACATGCCATTGCAGA 248
Db 377 cctgctcttttgcctgaccatgcctctatctgggcccgtctc-caaggaaaaag-gc-tggat 433
QY 249 CATCCTCTTTGTTCTTACTCTCCCACTTCTGGGCGAGTGAGTCATGCCACTGGTGGGTGGT 308
Db 434 ttccggcacgcccctgtgcaagggtgctcgtctgtggaaggaagcaactctacagtg 493
QY 309 TTTTACGACATGCCAGCTGCAAGTGTCTAAAGGCACTCTATGCCATCAACTTTAACTCGG 368
Db 494 aatcctgctctggcctgacagtgatgacccgactacccctggccattgtccatctact 551
QY 369 GATGCTGCTCTGACTTGCTAGCATGAGCCGGTACATGCCATTGTACAGCGGACT 426

RESULT 12
ID T79096 standard; cDNA; 1071 BP.
AC T79096;
DT 13-MAR-1998 (first entry)
DE Human CCR3 chemokine receptor coding sequence
KW CCR3 chemokine; mouse; primer; PCR; amplification; antagonist; human;
KW abnormal physiology; development; anti-viral; probe; hybridisation; ss.
OS Homo sapiens.
Key Location/Qualifiers
FT 1..1071
FT cds /*tag= a
FT /product= CCR3 chemokine receptor
PN W09721812-A2.
PD 19-JUN-1997.

W P E F H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Fri Nov 13 23:16:17 1998; MasPar time 1258.39 Seconds
Tabular output not generated.
1364.368 Million cell updates/sec
Title: >US-08-887-977-9
Description: (1-1119) from US08887977.seq
Perfect Score: 1119
N.A. Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCCCTCAA 1119
Comp: TACAAAGCTGAGTCACTT.....TCTTTCGACTCAGAGGGATT

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 1988617 seqs, 767163441 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: emb1-est55
1:em-est10 2:em-est11
Database: genbank-est107
3:gb-est1 4:gb-est10 5:gb-est11 6:gb-est12 7:gb-est13
8:gb-est14 9:gb-est15 10:gb-est16 11:gb-est17
12:gb-est18 13:gb-est19 14:gb-est2 15:gb-est20
16:gb-est21 17:gb-est22 18:gb-est23 19:gb-est24
20:gb-est25 21:gb-est26 22:gb-est27 23:gb-est28 24:gb-est29
25:gb-est30 26:gb-est31 27:gb-est32

Statistics: Mean 11.175; Variance 1.882; scale 5.936
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description	Pred. No.
No.								
C 1	160	14.3	461	18	AI045155	UI-R-C1-kk-c-08-0-UI.s		0.00e+00
2	133	11.9	312	13	AA790757	vv18g09.r1 Soares mous		3.37e-239
3	121	10.8	492	16	AA889777	a150h01.s1 Soares_NFL		6.10e-211
C 4	80	7.1	635	22	W28677	50b1 Human retina CDNA		5.87e-117
5	66	5.9	252	13	AA754459	97SN1787 Rice Immature		2.62e-86
C 6	57	5.1	252	13	AA754459	97SN1787 Rice Immature		2.62e-86
7	52	4.6	247	13	AA754458	97SN1784 Rice Immature		5.89e-57
8	50	4.5	130	13	AA750755	zab1f05.r1 Soares feta		5.89e-57
9	50	4.5	247	13	AA754458	97SN1784 Rice Immature		6.58e-53
C 10	50	4.5	247	13	AA754458	97SN1784 Rice Immature		6.58e-53
11	42	3.8	261	4	AA331167	EST35211 Embryo, 8 wee		3.31e-37
12	42	3.8	316	4	AA296456	EST11038 Umbilical vei		3.31e-37
13	42	3.8	466	14	R15256	yf89d03.r1 Homo sapien		3.31e-37

14	38	3.4	289	20	T28268	EST34714 Homo sapiens	1.02e-29
15	38	3.4	391	26	AA153412	mq66a06.r1 Soares 2NDM	1.02e-29
16	38	3.4	484	26	AA120018	mp93c11.r1 Soares 2NDM	1.02e-29
17	38	3.4	639	13	AA816049	vi14h02.r1 Barstead mo	1.02e-29
18	38	3.4	793	25	AA182270	mt83f01.r1 Soares mous	1.02e-29
19	36	3.2	539	27	FR0019844	F.rubripes GSS sequenc	4.36e-26
20	36	3.2	798	17	AI007060	ua81c03.r1 Soares mous	4.36e-26
21	36	3.2	816	15	AA205847	zq50c02.r1 Stratagene	4.36e-26
22	35	3.1	249	20	H29103	ym31f07.r1 Homo sapien	2.65e-24
23	35	3.1	2275	11	AF034173	Homo sapiens ntcon2 co	2.65e-24
24	34	3.0	166	16	AA915469	vz31c03.r1 Soares 2NDM	1.53e-22
25	34	3.0	342	10	AA621854	nq19g01.s1 NCI CGAP.Th	1.53e-22
26	34	3.0	382	14	R78657	vi74a05.r1 Homo sapien	1.53e-22
27	34	3.0	438	5	AA386601	EST99759 Pancreas tumo	1.53e-22
28	34	3.0	2275	11	AF034173	Homo sapiens ntcon2 co	1.53e-22
29	33	2.9	145	14	R23114	yh27b12.r1 Homo sapien	8.36e-21
30	32	2.9	241	8	AA566795	ZF-A179 zebrafish adul	4.31e-19
31	32	2.9	360	27	FR0024014	F.rubripes GSS sequenc	4.31e-19
32	32	2.9	370	8	AA225739	nc17b08.s1 NCI CGAP.Pr	8.36e-21
33	32	2.9	521	8	AA479467	zvl1f10.r1 Soares NHM	4.31e-19
34	31	2.8	386	21	AA5296	yzi1f09.s1 Homo sapien	2.08e-17
35	31	2.8	423	11	AA690396	vu52b02.r1 Soares mous	2.08e-17
36	31	2.8	424	5	AA035022	zk27c08.r1 Soares preg	2.08e-17
37	31	2.8	505	10	AA622439	nq40b08.s1 NCI CGAP.Co	2.08e-17
38	31	2.8	520	10	AA425767	zw47g03.s1 Soares tota	2.08e-17
39	31	2.8	544	5	AA148399	z144c09.s1 Soares preg	2.08e-17
40	31	2.8	551	17	AA862993	og99c06.s1 NCI CGAP.Ki	2.08e-17
41	31	2.8	552	5	AA148398	z144c09.r1 Soares preg	2.08e-17
42	31	2.8	607	24	AA058871	z196h03.s1 Stratagene	2.08e-17
43	31	2.8	782	25	AA203362	zx54h11.r1 Soares feta	2.08e-17
44	30	2.7	410	6	C05136	Human Heart CDNA, clon	9.42e-16
45	30	2.7	478	26	AA189997	mt98e07.r1 Soares mous	9.42e-16

ALIGNMENTS

1
LOCUS AI045155 461 bp mRNA EST 06-JUL-1998
DEFINITION UI-R-C1-kk-c-08-0-UI.s1 UI-R-C1 Rattus norvegicus CDNA clone
UI-R-C1-kk-c-08-0-UI 3', mRNA sequence.
ACCESSION AI045155
NID G3291974
KEYWORDS EST.
SOURCE Norway rat.
Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 461)
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence tag present in the CDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult Spleen library. CDNA Library Preparation: M. Fatima Bernaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics
Seq primer: M13 Forward.
Location/Qualifiers
1..461
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C1

library is a subtracted library derived from the UI-R-C0 library, which is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C1) was constructed as follows: PCR amplified cDNA inserts from UI-R-C0 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C0 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C1 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996).

```

/clone_ref="taxon:10116"
/clone="UI-R-Cl-kk-C-08-0-UI"
/clone_lib="UI-R-Cl"
/dev_stage="Adult"
/lab_host="PH108 (Life Technologies)"
100 a 113 c 108 g 138 t 2 others

```

Oy 724 AAAGCCATCCG 734

RESULT 3
 LOCUS AA889777 492 bp mRNA EST 03-APR-1998
 DEFINITION al50h01.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:1460785 3' similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE
 RECEPTOR TYPE 6 ; mRNA sequence.
 ACCESSION AA889777
 NID 93016656
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (Bases 1 to 492)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 109.

FEATURES
 source
 1. 492
 /organism="Homo sapiens"
 /note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NbHL19W, testis NHT, and B-cell
 NCL-CGAP-GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /clone="IMAGE:1460785"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /lab="119 c 117 g 122 t

BASE COUNT 134 a 119 c 117 g 122 t
 ORIGIN
 Query Match 10.8%; Score 121; DB 16; Length 492;
 Best Local Similarity 80.8%; Pred. No. 6.10e-211;
 Matches 172; Conservative 0; Mismatches 39; Indels 2; Gaps 2;
 Db 282 TTAGGGGGCCCCACCTTATCAATAGTGA-GGACGGCGCATCTCGTAACTCGCGT 340
 Cp 1119 TTAGGGAGACTCAGCTTCTATCATAGTGAAGGACGCGCATCTCGTATCTCGGT 1060
 Db 341 CTCATATGTTCCCGAAGAGGTTTCCGGGGAACCTCCGGGAAAGGCGCTGAGGA 400
 Cp 1059 CTCATCTGTTCCGGAGAATGTTTCTAGTAGTACCTCCCGGACAGGAGGCTGAGGA 1000
 Db 401 CTGGTACTTCTCCAGACACCAAGGTC-TTCAAGATCTTCCGAAAGTAGTTTCGAA 459
 Cp 999 CTGTACTTCTTCTCACACACACAGGTCCTTCAAGATCTTCAAGAAAGTAGTTTCGAA 940
 Db 460 CTTCTGGCCCAATAAAGGTTTAAGCACAGGGTT 492
 Cp 939 CTTCTGCCCAATAAAGCGTAGAGCACAGGTT 907

RESULT 4

LOCUS W28677 635 bp mRNA EST 08-MAY-1996
 DEFINITION 50b1 Human retina cDNA randomly primed sublibrary Homo sapiens
 cDNA, mRNA sequence.
 ACCESSION W28677
 NID g1308625
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (Bases 1 to 635)
 AUTHORS Macke, J., Smallwood, P. and Nathans, J.
 TITLE Adult Human Retina cDNA
 JOURNAL Unpublished (1996)
 COMMENT Contact: Dr. Jeremy Nathans
 Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
 Johns Hopkins School of Medicine
 725 North Wolfe Street, Baltimore, MD 21205
 Tel: 410 953 4678
 Fax: 410 614 0827
 Email: jeremy_nathans@mail.bs.jhu.edu
 Clones from this library are NOT available.
 PCR Primers
 FORWARD: CTTTTCAGCAAGTTCAGCTGGTTAAGT
 BACKWARD: GAGTGGCTTATGAGTATTTCTTCAGGGTAA
 Seq primer: GGGTAAAGCAAGAAATTT.
 Location/Qualifiers
 1. 635
 /organism="Homo sapiens"
 /note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI;
 Site_2: EcoRI; the library used for sequencing was a
 sublibrary derived from a human retina cDNA library.
 Inserts from retina cDNA library DNA were isolated,
 randomly primed, PCR amplified, size-selected, and cloned
 into lambda gt10. Individual plaques were arrayed and
 used as templates for PCR amplification, and these PCR
 products were used for sequencing."
 /db_xref="taxon:9606"
 /clone_lib="Human retina cDNA randomly primed sublibrary"
 /sex="mixed (males and females)"
 /tissue_type="retina"
 /dev_stage="adult"
 /lab_host="E. coli strain K802"
 /lab="136 c 143 g 128 t 94 others

BASE COUNT 134 a 136 c 143 g 128 t 94 others
 ORIGIN
 Query Match 7.1%; Score 80; DB 22; Length 635;
 Best Local Similarity 71.4%; Pred. No. 5.87e-117;
 Matches 177; Conservative 0; Mismatches 58; Indels 13; Gaps 13;
 Db 209 TTTCTNTACATNGGGGAGGANGAGGNATNN-CGATATCTGNGGCTCAGTGGTCT-CCG 266
 Cp 1104 TTTCTATCATAGTAGGAGGACGCGCATTCCTGTTATCTCGGCTCAGTGGTCTGCCG 1045
 Db 267 ANATATCTNN-CTCAGTACC-CCCGGAACAGAGAGANNNGANN-CTNGNTCT-CANNCT 322
 Cp 1044 AGAATGTTTCTGAGTACCTCCCGCACAGGAGAACCTGAGGACTGTACTTCTTCT 985
 Db 323 NANACACCACAGGTCCT-CAAGATCT-CAGAATGAGTCTTCTNN-CTNTCT-CCCA-TGAG 377
 Cp 984 CACACACCACAGGTCCTCAAGATCTTCAGAAAGTCTTCTGAACCTTCTGCCCAATAA 925
 Db 378 ANNGGGGTGCACAGNCTNAGNAGACNCAGACACN-CCAGNACN-CTNGACANTTNC 435
 Cp 924 AGCGTAGAGCAGGGTTCAGGCAGCAGTGCAGGAAAGCCAGGACTTCTGTGACAGTTT 865
 Db 436 CGTNTGC 443
 Cp 864 CGTATAGC 857

```

5
RESULT 5
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT
Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.osti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLr"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 5.9%; Score 66; DB 13; Length 252;
Best Local Similarity 10.3%; Pred. No. 2.62e-86;
Matches 24; Conservative 120; Mismatches 89; Indels 1; Gaps 1;
Db 18 WMTSTYBCHGNBVVVCVASHGNYMNVHNCBTRGTHCDCKNNV-WSTMWGTWVNVNVSVD 76
QY 453 ACTACCGCGCAGCAAAATCATCTCGCTTGTGTGGGGCTGCAGTCATCATCTCCAG 512
Db 77 WYWBVBNTKVDVGNHTRCSWRBVRMAHYHDYNCBBYNNNDYHMHBBMYBBTGCT 136
QY 513 CTAACATTTTGTCTTCAACCAAAATCAACACCAAGGAGCGAGCTGCTGTAACCCAA 572
Db 137 CTMWCWBHNTKCTASGWHSTNTYDKSSTNTWGTBYSYDKSMHGYWCS-BBVKYHTKVST 196
QY 573 GTACCAACTGCTCGGAGCCCATCAGTGGAGCTGCTGATGTGGGCTTGAGCTACT 632
Db 197 TRATRYTCVRKYCVMMMTKKVKKYHVBBGCHBTDSCKTKMTMKNHVMST 250
QY 633 CTTTGGTTTCTTATCCCTTGATGTTTCATGATATTTTGTACACGTTTCATTGT 686

6
RESULT 6
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT
Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.osti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLr"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 5.1%; Score 57; DB 13; Length 252;
Best Local Similarity 11.6%; Pred. No. 2.98e-67;
Matches 27; Conservative 117; Mismatches 86; Indels 3; Gaps 3;
Db 21 TSYBCHGNBVVVCVASHGNYMNVHNCBTRGTHCDCKNNVWSTMWGTWVNV-BNVSQDWHY 79
Cp 463 TCGCGGTAGTGTCTCTGGATCGGAGCGGAGTACTAGTCGCTGACATGGCGATCT 404
Db 80 WEVBNTKVDVGNHTRCSWRBVRMAHYHDYNCBBYNNNDYHMHBBMYBBTGCTM 139
Cp 403 ACCGTCCTCATGTAATCAAGTCAGGAGCAGCATCCCGAGTTAAAGTTGATGGCATAGA 344
Db 140 WCWBHNTKCTASGWHSTNTYDKSSTNTWGTBYSYDKSMHGYWCS-BBVKYHTKVSTTR 198
Cp 343 TGCCTTTAGCAACTGTCAGCTGGCATGCTGCTGAAACCCACGACGATGGCATCA 284
Db 199 ATRSYTCVRKYCVMMMTKKVKKYHVBBGCHBTDSCKTKMTMKNHVMSTT 251
Cp 283 CTGCCCAAGATGGGAGAGTAAGAACAAAGAG-GATGCTCTGCAATGCCCATGTT 232

7
RESULT 7
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

```

```

SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT
Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.osti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLr"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 5.1%; Score 57; DB 13; Length 252;
Best Local Similarity 11.6%; Pred. No. 2.98e-67;
Matches 27; Conservative 117; Mismatches 86; Indels 3; Gaps 3;
Db 21 TSYBCHGNBVVVCVASHGNYMNVHNCBTRGTHCDCKNNVWSTMWGTWVNV-BNVSQDWHY 79
Cp 463 TCGCGGTAGTGTCTCTGGATCGGAGCGGAGTACTAGTCGCTGACATGGCGATCT 404
Db 80 WEVBNTKVDVGNHTRCSWRBVRMAHYHDYNCBBYNNNDYHMHBBMYBBTGCTM 139
Cp 403 ACCGTCCTCATGTAATCAAGTCAGGAGCAGCATCCCGAGTTAAAGTTGATGGCATAGA 344
Db 140 WCWBHNTKCTASGWHSTNTYDKSSTNTWGTBYSYDKSMHGYWCS-BBVKYHTKVSTTR 198
Cp 343 TGCCTTTAGCAACTGTCAGCTGGCATGCTGCTGAAACCCACGACGATGGCATCA 284
Db 199 ATRSYTCVRKYCVMMMTKKVKKYHVBBGCHBTDSCKTKMTMKNHVMSTT 251
Cp 283 CTGCCCAAGATGGGAGAGTAAGAACAAAGAG-GATGCTCTGCAATGCCCATGTT 232

7
RESULT 7
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

```

Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.
 Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 Unpublished (1998)

Contact: Eun M.Y.

Department of Cytogenetics
 National Inst. of Agri. Sci. and Tech, RDA
 Suwon, Kyunggido, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307

Email: myeun@sun20.osti.re.kr
 Submitted by Baek Hie Namh, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bnaum@bioserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.

FEATURES

source

1. .247
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
 XhoI; Directional cDNA library inserted into lambda ZAPII
 vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"
 /clone="97SN1784"
 /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_host="E. coli SOLR"
 BASE COUNT 7 a 16 c 21 g 34 t 169 others
 ORIGIN

Query Match 4.6%; Score 52; DB 13; Length 247;
 Best Local Similarity 14.2%; Pred. No. 5.89e-57;
 Matches 32; Conservative 102; Mismatches 89; Indels 2; Gaps 2;

Db 5 TMNTVWRCGCCBAMWKKHTMMTBWCCVRRVGTNTNKGKTRTTWNDCSDNAHCY 64

QY 390 TAGCATGACGGGTACATCGCATGCTACAGGCGACTAGTATTCGGCTCGGATCCAG 449

Db 65 TVBMYARKSYGYBYYSWNVDTNTGTGVTGKTTNVHSGWNNRGSNSVYVYBAYCD 124

QY 450 AACACTACCGCGCAGCAAAATCATCTGCCTTGT-TGTGTGGGGGCGTGCAGTCATCATCT 508

Db 125 YBHTDRANHVDDTCTNDRGVCNTASDNGSATKRVGTGYKTDSDCGGCGWRKYTGS 184

QY 509 CCAGCTCAACTTTTG-TCTTCAACCAAAATACACACCCAGCGCATGCTGTGAA 567

Db 185 SBYBRGVNVVVRTTSMWTDKSKMBSMDMSRRSRVHYGRWBNK 229

QY 568 CCCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGCTG 612

RESULT 8
 LOCUS W04836 450 bp mRNA EST 23-APR-1996
 DEFINITION z881f05.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 298977
 5' similar to gb:D10925 C-C CHEMOKINE RECEPTOR TYPE 1 (HUMAN);,
 mRNA sequence.

ACCESSION W04836

NID g1277557

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;

1 (bases 1 to 450)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: mob.REGA+ET

High quality sequence stop: 402.

FEATURES

source

1. .450
 /organism="Homo sapiens"
 /note="Organ: lung; Vector: p7T3D (Pharmacia) with a
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 (5'-TGTACCAATCTGAAGTGGGAGCGCGCAATTTTTTTTTTTT-3'),
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NbHL19W."
 /db_xref="taxon:9606"
 /clone="298977"
 /clone_lib="Soares fetal lung NbHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 BASE COUNT 91 a 118 c 109 g 130 t 2 others
 ORIGIN

Query Match 4.6%; Score 52; DB 22; Length 450;
 Best Local Similarity 62.6%; Pred. No. 5.89e-57;
 Matches 129; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Db 58 TACTCCTTGTGTTATTTGCTGCTGGTGGAAACATCCTCGTGGTCTGCTTGTG 117

QY 130 TACTCCTTGTGTTATTTGCTGCTGGTGGAAATATCTGGTGGTATCACCTTGTCT 189

Db 118 CATACAAAGAGGCTAAACATGACCATCTACCTCTCTGATCAGATGAGGATGCTGAC 177

QY 190 TTTTATAAGAGGCGAGGTCTATGACAGAGCTATCTCTTGAACATGGCCATGCAGAC 249

Db 178 CTGCTCTCTCTGTCAGGCTCCCTTCTGGATCAGTACAAAGTTGAAGATGACTGGTT 237

QY 250 ATCCTCTTTTCTTACTCTCCCATCTGGGCGAGTGAATGCGACTGGTGGTGGT 309

Db 238 TTGGTGATGCCATGTGTAAGATCCT 263

QY 310 TTCAGCAATGCCAGTCCAAGTTCCT 335

RESULT 9

LOCUS AA790756 130 bp mRNA EST

DEFINITION w41808.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone

1244222 5' similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE RECEPTOR

TYPE 6 ; mRNA sequence.

ACCESSION AA790756

NID g2850876

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

1 (bases 1 to 130)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:657910

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .130
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCGGAATGGTGTGTGTGTGTGTGTGTGTGT
T 3'] double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
/db_xref="taxon:10090"
/clone="1244222"
/clone_lib="Soares mouse mammary gland NbMxg"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 35 a 35 c 27 g 33 t
ORIGIN

Query Match 4.5%; Score 50; DB 13; Length 130;
Best Local Similarity 73.9%; Pred. No. 6.58e-53;
Matches 85; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

Db 2 AAATCTTTCGGTGCCTCCAGAACTACGACAGTAAAGTCACTGTGTGGCAGTG 61
QY 427 AAGTCAATCCGGCTCCGATCCAGAACTACCGCGCAGCAAAATCATCTGCTTGTG 486

Db 62 TGGTTCATCTCCATCATCATCTCAAGCCCTACATTTAGTCTTCAACAGAAATAC 116
QY 487 TGGGGCTGTCAGTCATCATCTCAGCTCAACTTTT-GTCTTCAACCAAAATAC 540

RESULT 10
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
REFERENCE
AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,
Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,
Lee, M.C. and Eun, M.Y.
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)
Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeunesun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bnhahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers
1. .247
/organism="Oryza sativa"
/cultivar="Milvang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

Query Match 4.5%; Score 50; DB 13; Length 247;
Best Local Similarity 13.0%; Pred. No. 6.58e-53;
Matches 31; Conservative 108; Mismatches 97; Indels 2; Gaps 2;

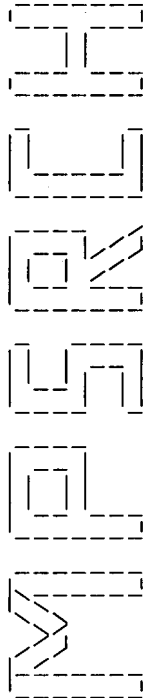
Db 1 HWDCTMTVWRCGCCBAAWNNKHTHMTBWCVRVGTGTTNKGKNGRTTTWDCSDNA 60
Cp 670 AAATAATCATGACATCAAGGATAAAGAACCAAGTAGTACTCAAGCCCCAACATCA 611

Db 61 HCR-YTVBYYARSKYGYGTBYYSNNVDNTGGTGYGKTYTVNHSGNNRCSNVYVWB 119
Cp 610 GCAGCTTCACCTGATGGCTCGACACAGTTTGGTACITGGTTTCACAGACATCGCTGC 551

Db 120 TAYCDYBYBDRANHVDDTRCNDRCYCNVTASDNGTSATKRVGYDKTSDCGGGGWRK 179
Cp 550 CTGGGTGTGTATTTTGGTTGAAGAAAAGTTGAGCTGGA-GATGATGACTGACAGC 492

Db 180 VYGSBBYBRGVNVVRYTTSMTDKSTKMBSDMRSRVHYGRWMBNKKRGMNRW 237
Cp 491 CCCACACAAAGGACAGATGATTTCGCGGCTAGTGTCTGTGATCGGACCGGA 434

RESULT 11
LOCUS AA331167 261 bp mRNA EST 21-APR-1997
DEFINITION ES135211 Embryo, 8 week 1 Homo sapiens cDNA 5' end similar to
vasoactive intestinal peptide receptor RDC1, mRNA sequence.
ACCESSION AA331167
NID 91983628
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 261)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.C., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitch, W.M., Fritchman, J.L., Geoghagen, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marnaros, S.M., Merrick, J.M.,
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.F., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sat Nov 14 00:36:39 1998; MasPar time 44.20 Seconds
Tabular output not generated. 1181.065 Million cell updates/sec

Title: >US-08-887-977-9
Description: (1-1119) from US08887977.seq
Perfect Score: 1119 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCTAA 1119
N.A. Sequence: Comp: TACAAAGCTGAGGTCACCTT.....TCTTCGACTCAGAGGATT

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 88822 seqs, 23323279 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: n-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1
Statistics: Mean 8.680; Variance 4.376; scale 1.984

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Query Match	Description	ID	
1	75	6.7	Sequence 18, Applicati	PCT-US93-1	2.18e-38
2	75	6.7	Sequence 6, Applicatio	PCT-US93-1	2.18e-38
3	75	6.7	Sequence 14, Applicati	PCT-US93-1	2.18e-38
4	75	6.7	Sequence 14, Applicati	PCT-US93-1	2.18e-38
5	61	5.5	Sequence 1, Applicatio	PCT-US93-1	9.02e-28
6	62	5.5	Sequence 23, Applicati	PCT-US93-1	1.61e-28
7	58	5.2	Sequence 14, Applicati	US-08-232	1.54e-25
8	57	5.1	Sequence 1, Applicatio	PCT-US93-0	8.47e-25
9	53	4.7	Sequence 1, Applicatio	PCT-US92-0	7.35e-22
10	51	4.6	Sequence 3, Applicatio	PCT-US95-0	2.10e-20
11	51	4.6	Sequence 1, Applicatio	PCT-US95-0	2.10e-20
12	49	4.4	Sequence 2, Applicatio	PCT-US95-0	5.83e-19
13	49	4.4	Sequence 3, Applicatio	PCT-US92-0	5.83e-19
14	49	4.4	Sequence 3, Applicatio	PCT-US95-0	5.83e-19
15	49	4.4	Sequence 2, Applicatio	US-08-202	5.83e-19
16	49	4.4	Sequence 1, Applicatio	US-08-418	5.83e-19
17	49	4.4	Sequence 1, Applicatio	US-08-410	5.83e-19
18	49	4.4	Sequence 1, Applicatio	US-08-076	5.83e-19
19	49	4.4	Sequence 1, Applicatio	PCT-US94-0	5.83e-19
20	49	4.4	Sequence 1, Applicatio	US-08-410	5.83e-19

21	48	4.3	1161	2	PCT-US93-1	Sequence 31, Applicati	3.05e-18
22	48	4.3	2254	2	PCT-US93-1	Sequence 27, Applicati	3.05e-18
23	45	4.0	1510	1	US-07-759	Sequence 4, Applicatio	4.16e-16
24	45	4.0	1748	1	US-08-202	Sequence 8, Applicatio	4.16e-16
25	43	3.8	1106	2	PCT-US92-0	Sequence 5, Applicatio	1.06e-14
26	43	3.8	1106	2	PCT-US95-0	Sequence 4, Applicatio	1.06e-14
27	43	3.8	1317	2	PCT-US93-1	Sequence 45, Applicati	1.06e-14
28	43	3.8	1737	1	US-08-076	Sequence 3, Applicatio	1.06e-14
29	43	3.8	1737	1	US-08-202	Sequence 4, Applicatio	1.06e-14
30	43	3.8	1737	2	PCT-US94-0	Sequence 2, Applicatio	1.06e-14
31	40	3.6	652	2	PCT-US92-0	Sequence 2, Applicatio	1.28e-12
32	40	3.6	1244	1	US-07-816	Sequence 7, Applicatio	1.28e-12
33	39	3.5	1080	2	PCT-US95-1	Sequence 3, Applicatio	6.23e-12
34	36	3.2	1572	1	US-08-041	Sequence 5, Applicatio	6.88e-10
35	36	3.2	1572	1	US-08-417	Sequence 5, Applicatio	6.88e-10
36	36	3.2	1679	2	PCT-US93-1	Sequence 6, Applicatio	6.88e-10
37	36	3.2	1679	2	PCT-US94-0	Sequence 3, Applicatio	6.88e-10
38	36	3.2	1679	1	US-08-076	Sequence 5, Applicatio	6.88e-10
39	32	2.9	1351	1	US-07-816	Sequence 5, Applicatio	2.86e-07
40	29	2.6	69	2	PCT-US93-1	Sequence 1, Applicatio	2.30e-05
41	29	2.6	215	1	US-08-238	Sequence 5, Applicatio	2.30e-05
42	29	2.6	720	2	PCT-US93-1	Sequence 35, Applicati	2.30e-05
43	28	2.5	74	2	PCT-US95-1	Sequence 100, Applicati	9.59e-05
44	28	2.5	105	1	US-07-865	Sequence 13, Applicati	9.59e-05
45	28	2.5	1842	1	US-08-446	Sequence 1, Applicatio	9.59e-05

ALIGNMENTS

RESULT 1
ID PCT-US93-11153-18 STANDARD; DNA; UNC; 1900 BP.
AC xxxxxx
DE Sequence 18, Application PC/TUS9311153
CC Sequence 18, Application PC/TUS9311153
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11153
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1900 base pairs
CC TYPE: nucleic acid

CC ADDRESSEE: Townsend and Townsend Khourie and Crew
CC STREET: One Market Plaza, Steuart Tower, Suite 2000
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94610
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/012,988A
CC FILING DATE: 19930128
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Weber, Kenneth A.
CC REGISTRATION NUMBER: 31,677
CC REFERENCE/DOCKET NUMBER: 15280-118
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-543-9600
CC TELEFAX: 415-543-5043
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2156 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: primer_bind
CC LOCATION: 259..275
CC FEATURE:
CC NAME/KEY: primer_bind
CC LOCATION: complement (868..884)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 63..1128
CC SEQUENCE 2156 BP; 538 A; 547 C; 504 G; 567 T; 0 OTHER.

Query Match 5.5%; Score 61; DB 1; Length 2156;
Best Local Similarity 60.4%; Pred.No. 9.02e-28;
Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Db 183 TACTCCTTGATTTGTCATGGCTGCTGGAAACATCCCTGGTCTGCTGCTTGTG 242
QY 130 TACTCCTTGATTTGTCATGGCTGCTGGGAATATCTGGTGGTATCACCTTTGCT 189
Db 243 CAATACAGAGGCTAAACACATGACAGCATCTACCTCTGAAACCTGGCCATTTCTGAC 302
QY 190 TTTTATAAGAGGCGCAGGTCTATGACAGACGCTCTATCTTGAACATGGCCATTCGAC 249
Db 303 CTGCTCTCTGTTACGCTTCCCTCTGGATCGACTACAAAGTTGAAGATGACTGGGTT 362
QY 250 ATCCTCTTTGTTTCTTACTCTCCCACTTGGGCGAGTATGACCATGGTGGCGGTT 309
Db 363 TTTGGGTATGCCATGTGTAAGATCCTCTCTGGGTTTATTACAGAGGCTTGTACAGCGAG 422
QY 310 TTCAGCAATGCCAGTGCAGTTGCTTAAAGGCACTATGCGCATCAACTTTAACTGCGGG 369
Db 423 ATCTTTTCAATCCTCTGCTGACGATTGACAGGTAACCTGGCCATCGTCCAGC 475
QY 370 ATGCTGCTCTGACTTGCATTAGCATGACCGGTATACCGCATTCGCGCATTTGTACGSC 422

RESULT 6

ID PCT-US93-11153-23 STANDARD; DNA; UNC; 2751 BP.
AC xxxxxx

DE Sequence 23, Application PC/TUS93111153
CC Sequence 23, Application PC/TUS93111153
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald

CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schwellkart, Vicki L.
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11153
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA: US 07/977,452
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2751 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: intron
CC LOCATION: 1..691
CC FEATURE:
CC NAME/KEY: exon
CC LOCATION: 692..1771
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 692..1768
CC FEATURE:
CC NAME/KEY: polyA_signal
CC LOCATION: 2341..2348
CC SEQUENCE 2751 BP; 651 A; 747 C; 684 G; 669 T; 0 OTHER.
Query Match 5.5%; Score 62; DB 2; Length 2751;
Best Local Similarity 63.1%; Pred.No. 1.61e-28;
Matches 185; Conservative 0; Mismatches 105; Indels 3; Gaps 2;
Db 824 TATTCTGTCATCTGCTCGTGGCCCTGCTCGCAACGGGCTGGTGATCTAGCTACATC 883
QY 130 TACTCCTTGATTTGTTCTTGGCCCTCTGGGGAATATTCTGGTGGTATCACCTTTGCT 189
Db 884 TATTCAAGAGGCTCAAGACCATGACGGATACCTACCTGCTCAACCTGGCCGTCGACAG 943
QY 190 TTTTATAAGAGGCGCAGGTCTATGACAGACGCTATCTCTTGACATGGCCATTCGAC 249
Db 944 ATCCTTTTCTCTTAATTTCTCCCTTTCTGGGCTACAGGCAAGCCAA--GTCC-TGATC 1000
QY 250 ATCCTCTTTGTTTCTTACTCTCCCACTTGGGCGAGTGAAGTATGCTGCTGCTGGGTT 309
Db 1001 TTTGGCGTCTACCTGTTGAAGGCATCTTTGGCATCTATAAGTTAGCTTCTTCACGGG 1060
QY 310 TTCAGCAATGCCAGGTCGAAGTTGCTTAAAGGCATCTATGCCATCACTTTAACTCGGG 369

CC LOCATION: 81..1160
SQ SEQUENCE 1979 BP; 530 A; 435 C; 451 G; 563 T; 0 OTHER.

Query Match 4.6%; Score 51; DB 2; Length 1979;
Best Local Similarity 61.2%; Pred. No. 2.10e-20;
Matches 180; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

Db 224 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 283
QY 129 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 188
Db 284 AAATGCAAAAGCTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 343
QY 189 TTTTATGAAGAGCCAGCTCTATGACAGACGCTCTCTCTTGAACATGSCCATTCGAGA 248
Db 344 TCTGCTTTTCTATATCTCTCCCATTTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 400
QY 249 CATCTCTTTGTTTCTTACTCTCCCATTTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 308
Db 401 CTTTGGGAATGCAATGTGCAAAATTATTCACAGGCTGTATCATCATCGGTTATTTGGCGG 460
QY 309 TTTGAGCAATGCAATGTGCAAAATTATTCACAGGCTGTATCATCATCGGTTATTTGGCGG 368
Db 461 AATCTCTTCTCATCATCTCTCTGACATGATAGATACCTGGCTGATTTGTCATGCG 514
QY 369 GATGCTGCTCTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 422

RESULT 11
ID PCT-US95-00476-1 STANDARD; DNA; UNC; 2232 BP.
AC xxxxxx

Sequence 1, Application PC/TUS9500476
Sequence 1, Application PC/TUS9500476
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
TITLE OF INVENTION: PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson
STREET: 201 N. Figueroa Street, 5th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00476
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berliner, Robert
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-291
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-977-1001
TELEFAX: 310-977-1003
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2232 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

CC NAME/KEY: CDS
CC LOCATION: 40..1161
SQ SEQUENCE 2232 BP; 602 A; 464 C; 508 G; 658 T; 0 OTHER.

Query Match 4.6%; Score 51; DB 2; Length 2232;
Best Local Similarity 61.2%; Pred. No. 2.10e-20;
Matches 180; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

Db 183 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 242
QY 129 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 188
Db 243 AAATGCAAAAGCTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 302
QY 189 TTTTATGAAGAGCCAGGCTCTATGACAGACGCTCTCTCTTGAACATGSCCATTCGAGA 248
Db 303 TCTGCTTTTCTTATTAATCTCTCCCATTTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 359
QY 249 CATCTCTTTGTTTCTTACTCTCCCATTTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 308
Db 360 CTTTGGGAATGCAATGTGCAAAATTATTCACAGGCTGTATCATCATCGGTTATTTGGCGG 419
QY 309 TTTGAGCAATGCAATGTGCAAAATTATTCACAGGCTGTATCATCATCGGTTATTTGGCGG 368
Db 420 AATCTCTTCTCATCATCTCTCTGACATGATAGATACCTGGCTGATTTGTCATGCG 473
QY 369 GATGCTGCTCTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 422

RESULT 12
ID PCT-US95-03032-2 STANDARD; DNA; UNC; 1176 BP.
AC xxxxxx

Sequence 2, Application PC/TUS9503032
Sequence 2, Application PC/TUS9503032
GENERAL INFORMATION:
APPLICANT: Repligen Corporation
APPLICANT: the Trustees of Boston University
TITLE OF INVENTION: ANTIBODIES TO INTERLEUKIN-8 RECEPTORS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: WordPerfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03032
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/237,937
FILING DATE: 02-MAY-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/210,250
FILING DATE: 15-MAR-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,842
FILING DATE: 09-DEC-91
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/726,606
FILING DATE: 09-JUL-91
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/685,101
FILING DATE: 10-APR-91
ATTORNEY/AGENT INFORMATION:

WIREH

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Nov 14 00:39:25 1998; MasPar time 901.46 Seconds
Tabular output not generated. 1314.048 Million cell updates/sec

Title: >US-08-887-977-9
Description: (1-1119) from US08887977.seq
Perfect Score: 1119

N.A. Sequence: 1 ATGTTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119
Comp: TACAAAGCTGAGTCACTT.....TCTTCGACTCAGAGGATT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1639711 seqs, 529296662 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU7
29:NEWU8 30:NEWU9

Statistics: Mean 10.319; Variance 3.340; scale 3.090

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1119	100.0	1119	19	US-08-675- Sequence 7, Applicatio	0.00e+00
2	1119	100.0	1119	21	US-08-887- Sequence 9, Applicatio	0.00e+00
3	1104	98.7	1255	23	US-09-023- Sequence 993, Applicat	0.00e+00
4	1104	98.7	1679	23	US-09-016- Sequence 1097, Applicat	0.00e+00
5	1070	95.6	1137	23	US-09-023- Sequence 970, Applicat	0.00e+00
6	246	22.0	375	20	US-08-706- Sequence 3273, Applicat	4.84e-269
C 7	96	8.6	578	19	US-08-624- Sequence 14, Applicati	2.29e-80
C 8	96	8.6	578	18	US-08-569- Sequence 14, Applicati	2.29e-80
C 9	76	6.8	383	16	US-08-446- Sequence 8, Applicatio	1.19e-56
C 10	76	6.8	383	16	US-08-446- Sequence 8, Applicatio	1.19e-56
C 11	76	6.8	383	16	US-08-446- Sequence 8, Applicatio	1.19e-56
C 12	76	6.8	383	17	US-08-311- Sequence 8, Applicatio	1.19e-56
C 13	76	6.8	383	17	US-08-451- Sequence 8, Applicatio	1.19e-56
C 14	76	6.8	383	16	US-08-446- Sequence 8, Applicatio	1.19e-56
15	75	6.7	1500	23	US-09-023- Sequence 1465, Applicat	1.19e-56
16	75	6.7	1900	14	US-08-245- Sequence 18, Applicati	1.74e-55

17	75	6.7	2058	11	US-07-977- Sequence 6, Applicatio	1.74e-55
18	75	6.7	2058	14	US-08-245- Sequence 6, Applicatio	1.74e-55
19	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
20	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
21	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
22	75	6.7	2154	15	US-08-352- Sequence 1, Applicatio	1.74e-55
23	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
24	75	6.7	2154	23	US-09-023- Sequence 1094, Applicat	1.74e-55
25	75	6.7	2154	23	US-09-016- Sequence 1214, Applicat	1.74e-55
26	75	6.7	2160	14	US-08-245- Sequence 14, Applicati	1.74e-55
27	73	6.5	578	18	US-08-569- Sequence 14, Applicati	3.66e-53
28	73	6.5	578	19	US-08-624- Sequence 9, Applicati	3.66e-53
29	65	5.8	340	19	US-08-617- Sequence 9, Applicatio	5.61e-44
30	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
31	65	5.8	383	15	US-08-311- Sequence 8, Applicatio	5.61e-44
32	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
33	65	5.8	383	17	US-08-451- Sequence 8, Applicatio	5.61e-44
34	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
35	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
36	64	5.7	2577	23	US-09-016- Sequence 1095, Applicat	7.66e-43
37	63	5.6	339	19	US-08-617- Sequence 8, Applicatio	1.04e-41
38	61	5.5	1094	22	US-08-902- Sequence 1, Applicatio	1.86e-39
39	61	5.5	1495	23	US-09-023- Sequence 1021, Applicat	1.86e-39
40	61	5.5	1495	1	PCT-US93-1 Sequence 8, Applicatio	1.86e-39
41	61	5.5	1495	15	US-08-308- Sequence 8, Applicatio	1.86e-39
42	61	5.5	1495	23	US-09-016- Sequence 1190, Applicat	1.86e-39
43	62	5.5	2085	14	US-08-245- Sequence 65, Applicati	1.39e-40
44	61	5.5	2156	23	US-09-023- Sequence 1247, Applicat	1.86e-39
45	62	5.5	2751	14	US-08-245- Sequence 23, Applicati	1.39e-40

ALIGNMENTS

RESULT 1
ID US-08-675-814-7 STANDARD; DNA; UNC; 1119 BP.
AC xxxxxx
DT
DE Sequence 7, Application US/08675814
CC Sequence 7, Application US/08675814
CC GENERAL INFORMATION:
CC APPLICANT: Gish, Kurt C.
CC APPLICANT: Schall, Thomas J.
CC APPLICANT: Vicari, Alain
CC APPLICANT: Zlotnick, Albert
CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DNAX Research Institute
CC STREET: 901 California Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304-1104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA: US/08/675.814
CC APPLICATION NUMBER: US/08/675.814
CC FILING DATE: 05-JUL-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ching, Edwin P. 34,090
CC REGISTRATION NUMBER: 34,090
CC REFERENCE/DOCKET NUMBER: DX0589
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-852-9196
CC TELEFAX: 415-496-1200
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1119 base pairs
CC TYPE: nucleic acid

Db	268	ACCTTTGCTTTTATTAAGAAGGCCAGGTCTATGACAGACGCTCTATCTTTGAACATGGCC	327
Qy	181	ACCTTTGCTTTTATTAAGAAGGCCAGGTCTATGACAGACGCTCTATCTTTGAACATGGCC	240
Db	328	ATTGACAGACATCCTCTTTGTTCTTACTCTCCCAATCTGGCAGTGAGTCATGCCACTGGT	387
Qy	241	ATTGACAGACATCCTCTTTGTTCTTACTCTCCCAATCTGGCAGTGAGTCATGCCACTGGT	300
Db	388	GGTGGGTTTTTCAGCAATGCCAGTGCACAAAGTTCCTAAAAGGCATCTATGGCATCAACTTT	447
Qy	301	GGTGGGTTTTTCAGCAATGCCAGTGCACAAAGTTCCTAAAAGGCATCTATGGCATCAACTTT	360
Db	448	AACTCGGGAGTGTGCTCTGACTTCGATTAGCATGGACGGGTACATCGCCATGTACAG	507
Qy	361	AACTCGGGAGTGTGCTCTGACTTCGATTAGCATGGACGGGTACATCGCCATGTACAG	420
Db	508	GGGACTAAGTCATTCGGGTCCGATCCAGAACACTACCGGCGACGACAAATCATCTGCCTT	567
Qy	421	GGGACTAAGTCATTCGGGTCCGATCCAGAACACTACCGGCGACGACAAATCATCTGCCTT	480
Db	568	GTGTGTGGGGGTGTCAAGTCATCATCTCCAGCTCAACTTTTCTCTTCAACCAAAAATAC	627
Qy	481	GTGTGTGGGGGTGTCAAGTCATCATCTCCAGCTCAACTTTTCTCTTCAACCAAAAATAC	540
Db	628	AAACCCAAAGCAGCGATGTCTGTGAACCCAAAGTACCAAGCTCTCGGAGCCCATCAGG	687
Qy	541	AAACCCAAAGCAGCGATGTCTGTGAACCCAAAGTACCAAACTCTCGGAGCCCATCAGG	600
Db	688	TGGAAGCTGCTCATGTTGGGGCTTGAGCTACCTTTTGGTTTCTTTATCCCTTTGATGTTT	747
Qy	601	TGGAAGCTGCTCATGTTGGGGCTTGAGCTACCTTTTGGTTTCTTTATCCCTTTGATGTTT	660
Db	748	ATGATATTTTGTATACAGTTCATTGTCAAAACCTTGGTGCAAGCTCAGAAATCTAAAAAG	807
Qy	661	ATGATATTTTGTATACAGTTCATTGTCAAAACCTTGGTGCAAGCTCAGAAATCTAAAAAG	720
Db	808	CACAAAGCCATCCGTGTAAATCATAGCTGTGGTCTTTGTTTCTTGCTTGTCAGATTCCT	867
Qy	721	CACAAAGCCATCCGTGTAAATCATAGCTGTGGTCTTTGTTTCTTGCTTGTCAGATTCCT	780
Db	868	CATAACATGGTCTGCTTCTGACGGTGCATAATTCGGGTAAATGAACCGATCCTGCCAG	927
Qy	781	CATAACATGGTCTGCTTCTGACGGTGCATAATTCGGGTAAATGAACCGATCCTGCCAG	840
Db	928	AGCGAAAAGCTAATTGGCTATACGAAAACCTGTACAGAAAGTCTCGCTTTCTTCGCACTGC	987
Qy	841	AGCGAAAAGCTAATTGGCTATACGAAAACCTGTACAGAAAGTCTCGCTTTCTTCGCACTGC	900
Db	988	TGCTTGAACCCCTGTGCTCTACGCTTTTATTTGGCAGAGTTCAGAAACACTTCTTGAAG	1047
Qy	901	TGCTTGAACCCCTGTGCTCTACGCTTTTATTTGGCAGAGTTCAGAAACACTTCTTGAAG	960
Db	1048	ATCTTGAAGGACCTGTGGTGTGTGAGAAGGAAGTACAAGTCTTCAGGCTCTCCTGTGCC	1107
Qy	961	ATCTTGAAGGACCTGTGGTGTGTGAGAAGGAAGTACAAGTCTTCAGGCTCTCCTGTGCC	1020
Db	1108	GGGAGGTACTCAGAAAACATTTCTCGGCAGACAGCTGAGACCGCAGATAACGACAATGCG	1167
Qy	1021	GGGAGGTACTCAGAAAACATTTCTCGGCAGACAGCTGAGACCGCAGATAACGACAATGCG	1080
Db	1168	TCGTCTTCACTATGTAGAAAGCTGAGTCTCCCTAA	1206
Qy	1081	TCGTCTTCACTATGTAGAAAGCTGAGTCTCCCTAA	1119

RESULTS

ID	US-09-016-434-1097	STANDARD	DNA	UNC	1679 BP
AL00001					

AC
XXXXXX

DT

Sequence 1097 Application PS/00015434

sequence 1097, Application US/09016434

Sequence 1097, Application

CC GENERAL INFORMATION:

APPLICANT: Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 TYPE OF INVENTION: PATHWAY GENE EXPRESSION
 NUMBER OF SEQUENCES: 1490
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/016,434
 FILING DATE: HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0002 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 1097:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1679 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GENBANK
 CLONE: g1245058
 SEQUENCE 1679 BP: 404 A: 385 C: 381 G: 509 T: 0 OTHER.

Query Match 98 78. score 1104. DB 23. Length 1670.

Query Match

Best Local Similarity 99.6%; Pred. No. 0.00e+00;

db 256 ATGTTTCCACTCCAGTCAAGTCAATATTGTCTCAGTCAATACCTCATATTATCAGCTTC 315

DB 238 A T G T T T C G A C T C C A G T G A A G A T T A T T T G T C A A T A C T C A T A T A C T C A G T T G 313

THE

Db 316 ATTCTGAGATGTTACTGTGCTCCTTGCAGGAGGTCCAGGGCTATTGTGTAC 375

QY 61 ATCTGAGATGTTACTGTGCTCCTTG CAGGAGGTCAGGCAGTTCTCCAGGCTATTGTAC 120

QY 121 CGAATTGCCTACTCCTTGATCTGTGTTCTTGGCCTCCTGGGAATATCTGGTGGTGATC 180

DB 435 ACCCTTGCCTTTATTAAGAAGGCCAGGTCTATGACAGACGCTATCTCTTGAAACAAGGCC 494

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1991	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100										

Db 495 ATGCAGACATCCTCTTTGTTCTTACTCTCCCATTTCTGGGCAGTGAGTCAATGCCACTGGT 554

DD 433 ATTCAGACATCCCTCTTGTCTTACCTCCATCTGGGACGTGAGTCATGCCACCTGGT 334

Db 555 GCGTGGGTTTTCAGCAATGCCACGTGCAAGTTGCTAAAGGCATCTATGCCATCAACTTT 614

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

Query Match 8.6%; Score 96; DB 19; Length 578;
Best Local Similarity 12.7%; Pred. No. 2,298-80;
Matches 72; Conservative 264; Mismatches 226; Indels 7; Gaps 7;
Db 3 SARVTSRDSGRVGNMNVSKTSSTSSDKYKAHSGSHRSVTVKASGKMKYDVSGKKSKD 62
Cp 792 GACCATGTTATGAGCAATCTGACAGCCAGAAACACACAGCAGCTATGA-TTACAC 734
Db 63 GSDYSHKSKKNVTVKRYVADSKRGCMKNKHS-DMTGTWNNMNAKTSWSGKGT 121
Cp 733 GGATGGCTTTGCTTTAGCAATCTGAGCTTGACCAAGGTTTTCACAAATGAACGTGT 674
Db 122 RDDSADYHDYGTDSKWKHSTSVTKVAHTWNRVAKHNRSHCTDNVKTGANTGNKGAUG 181
Cp 673 AACAAATATCATGAACATCAAGGATTAAGAAACCAAGAGTAGCTCAAGCCCAACA 614
Db 182 VSMNGTSGVNSHTSGSK-KRRNNYMNRAKDKNSTHRTHWGDNRYDVTWATKYADSHDT 240
Cp 613 TCAGCAGCTTCCACCTGATGGCTCGGACACAGTTTGGTACTTGGGTTACACACATCGC 554
Db 241 RRKVTATYRTRDKYATKTKRYVADSKRGCMKNKHS-DMTGTWNNMNAKTSWSGKGT 300
Cp 553 TGCCCTGGCTTTGCTTTAGCAATCTGAGCTTGACCAAGGTTTTCACAAATGAACGTGT 494
Db 301 GVTSVKNGGTVDGRGYATKTKRYHSCVSKSGVNGSVYKTKSDYDHS-KSSDSD 359
Cp 493 GCCCCACACACAAAGGAGATGTTTCTGCGCGGTAGTGTCTGGATCGGAGCGGA 434
Db 360 SYGCGARATYTHHGTGKTSKRYKVDKTRDSSGKTSKSHDMKWVTSRACSGST 419
Cp 433 ATGACTTAGTCCCTGTAAGTGGCATGACCGGTCATGCTAAGTGAAGTGAAGTGAAGTGA 374
Db 420 NNYMGVGMHYKTSDFATSKYDSDGCRGSGTSGKSTANRGRTSRSDGKNAGDTRKNYG 479
Cp 373 GCATCCCGCAGTTAAAGT-TGATGGCATAGATGCCCTTTTACCAACTGACGTTGCGATTG 315
Db 480 SSSKARKDSKMRKCGSSVTKAADRGCGKVARRSCTSSSAGRAAGGDKSGKKTSSAVAKR 539
Cp 314 CTGAAACCCACGACAGTGGCATGA-CTCACTGCCAGATGGGAGAGTAAGAACAA 256
Db 540 KRSNTTTRVKSVAHHSKGRDYRDNTNHHG 568
Cp 255 GAGGATGT-CTGCAATGGCCATGTTCAAG 228

RESULT 8
ID US-08-569-578-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT
DE
Sequence 14, Application US/08569578
CC Sequence 14, Application US/08569578
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: IDENTIFICATION, CLONING, AND USE OF
CC TITLE OF INVENTION: SHC-ASSOCIATED INOSITOL POLYPHOSPHATE 5-PHOSPHATASE
CC TITLE OF INVENTION: (SAPTASE)
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION

CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/569,578
CC FILING DATE: 08-DEC-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-601-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.
Query Match 8.6%; Score 96; DB 18; Length 578;
Best Local Similarity 12.7%; Pred. No. 2,298-80;
Matches 72; Conservative 264; Mismatches 226; Indels 7; Gaps 7;
Db 3 SARVTSRDSGRVGNMNVSKTSSTSSDKYKAHSGSHRSVTVKASGKMKYDVSGKKSKD 62
Cp 792 GACCATGTTATGAGCAATCTGACAGCCAGAAACACACAGCAGCTATGA-TTACAC 734
Db 63 GSDYSHKSKKNVTVKRYVADSKRGCMKNKHS-DMTGTWNNMNAKTSWSGKGT 121
Cp 733 GGATGGCTTTGCTTTAGCAATCTGAGCTTGACCAAGGTTTTCACAAATGAACGTGT 674
Db 122 RDDSADYHDYGTDSKWKHSTSVTKVAHTWNRVAKHNRSHCTDNVKTGANTGNKGAUG 181
Cp 673 AACAAATATCATGAACATCAAGGATTAAGAAACCAAGAGTAGCTCAAGCCCAACA 614
Db 182 VSMNGTSGVNSHTSGSK-KRRNNYMNRAKDKNSTHRTHWGDNRYDVTWATKYADSHDT 240
Cp 613 TCAGCAGCTTCCACCTGATGGCTCGGACACAGTTTGGTACTTGGGTTACACACATCGC 554
Db 241 RRKVTATYRTRDKYATKTKRYVADSKRGCMKNKHS-DMTGTWNNMNAKTSWSGKGT 300
Cp 553 TGCCCTGGCTTTGCTTTAGCAATCTGAGCTTGACCAAGGTTTTCACAAATGAACGTGT 494
Db 301 GVTSVKNGGTVDGRGYATKTKRYHSCVSKSGVNGSVYKTKSDYDHS-KSSDSD 359
Cp 493 GCCCCACACACAAAGGAGATGTTTCTGCGCGGTAGTGTCTGGATCGGAGCGGA 434
Db 360 SYGCGARATYTHHGTGKTSKRYKVDKTRDSSGKTSKSHDMKWVTSRACSGST 419
Cp 433 ATGACTTAGTCCCTGTAAGTGGCATGACCGGTCATGCTAAGTGAAGTGAAGTGAAGTGA 374
Db 420 NNYMGVGMHYKTSDFATSKYDSDGCRGSGTSGKSTANRGRTSRSDGKNAGDTRKNYG 479
Cp 373 GCATCCCGCAGTTAAAGT-TGATGGCATAGATGCCCTTTTACCAACTGACGTTGCGATTG 315
Db 480 SSSKARKDSKMRKCGSSVTKAADRGCGKVARRSCTSSSAGRAAGGDKSGKKTSSAVAKR 539
Cp 314 CTGAAACCCACGACAGTGGCATGA-CTCACTGCCAGATGGGAGAGTAAGAACAA 256
Db 540 KRSNTTTRVKSVAHHSKGRDYRDNTNHHG 568
Cp 255 GAGGATGT-CTGCAATGGCCATGTTCAAG 228

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,937
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH191-09FW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 383 BP; 27 A; 25 C; 26 G; 34 T; 271 OTHER.

Query Match 6.8%; Score 76; DB 16; Length 383;
Best Local Similarity 11.1%; Pred. No. 1.19e-56;
Matches 40; Conservative 172; Mismatches 143; Indels 4; Gaps 4;

Db 16 HWTRASTHVSVNNDVMDNNGAVCKDCYRSTCDNKSCMSNCSTCKVCVAVVRKNDN 75
Cp 487 ACACAACAGCAGAGATGATTTCTCGCGGTAGTGTCTGGATCGGAGCGG-GAATGAC 429
Db 76 TTVCHDKYHDDAASKMKKKGKMTSCSSDNDNSYNTSNDVYTGSGVASVYCYVRNRKS 135
Cp 428 TTAGTCGCTGTACATGGATGGATGTACCGTGCCATGTAATGCAAGTCAGGAGCAGATC 369
Db 136 STWTGTRKMSHCADRSSTSCANNHNHTDVGKRAVYKAKKNTSTVAVKYDYHYSWK 195
Cp 368 CGCAGTTAAAGTTGATGCATAGATGCCTTTTAGCAACTT-GCACGTGGCATTCGTGAA 310
Db 196 DRKDSNKNTARKTKYKTAHAKNYTRHVSWDNRNVGVSARGSHSDHTCGRKMVRHK 255
Cp 309 AACCACGACCCAGTGGCATGCTACTGCCAGATGGGAGAGTAAGAACAAGAGGAT 250
Db 256 SSVNNDTCCCGSRGYSVDDANSVGTARYNAV-SRMNNASKTDVYSNAVWMTSRCNA 314
Cp 249 GTCTGCAATGGCCATGTTCAAGAGATAGAGCTGTGTCATAGACCTGGCCTTCTTATAAA 190
Db 315 VGVKDYGSKVRDVVSMKNDVDRGRNSWNHGMVCTTCDWDHARTACVARSHDRSGRS 373
Cp 189 ACAGAGGTGATCACC-ACCAGATATTTCCAGGAGGCCAAGAACACAGATCAAGAG 132

RESULT 15
ID US-09-023-655-1465 STANDARD; DNA; UNC; 1500 BP.
AC xxxxxx
DE
DT Sequence 1465, Application US/09023655
CC Sequence 1465, Application US/09023655
CC GENERAL INFORMATION:
CC APPLICANT: Cocks, Benjamin G.
CC APPLICANT: Susan G. Stuart
CC APPLICANT: Jeffrey J. Seilhamer
CC TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
CC TITLE OF INVENTION: EXPRESSION
CC NUMBER OF SEQUENCES: 1508
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE

CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/023,655
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Zeller, Karen J.
CC REGISTRATION NUMBER: 37,071
CC REFERENCE/DOCKET NUMBER: PA-0001 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 1465:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1500 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GENBANK
CC CLONE: g673391
CC SEQUENCE 1500 BP; 324 A; 457 C; 363 G; 356 T; 0 OTHER.

Query Match 6.7%; Score 75; DB 23; Length 1500;
Best Local Similarity 63.7%; Pred. No. 1.74e-55;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 483 TGTGTCCTCAAGAAGACGCTGGGGAACCTT-TAAAGCTGTGTTCTCCCTATCATGTACTCC 541
Qy 76 TGTGCTCTTCGAGGAGTCAGGAGTTCCTCAGGCTATTTGTACCGAATTCCTACTCC 135
Db 542 ATCAATTTCTTGTGGGCTACTGGGCAATGGCTGGTGGTGTGACCTATATCTATTC 601
Qy 136 TTGATCTGTGTTGTGGCTCTCTGGGAATATCTGTTGGTGATCACCTTGTCTTTAT 195
Db 602 AAGAGGCTCAAGACCATGACGATACCTGCTCAACCTGGCGGTGGCAGACATCCTC 661
Qy 196 AAGAAGGCCAGGTCTATGACAGAGCTCTATCTTTGAACATGGCCATTCGACATCCTC 255
Db 662 TTCTCTCTGACCTTCCCTCTCTGGGCTTACAGCGGCGCAA--GTCC-TGGGCTTTCGGT 718
Qy 256 TTTGTTCTTACTCTCCCAATTCGAGCACTGAGTATGCACTGGTGGTGGTGGTTCAGC 315
Db 719 GTCCACTTTTGAAGCTCATCTTTGCCATCTACAGATGAGCTTCTTCTAGTGGCATGCTC 778
Qy 316 AATGCCAGTGCAGTTGCTTAAAGGCATCTATGCCATCAACTTAACTGCGGGATGCTG 375
Db 779 CTACTTCTTTGCATCAGCATTTGACCGCTACGTGGCCATCGTCCAGGC 825
Qy 376 CTCCTGACTTGCATTAGCATGGACCGGTACATCGCAATGTACAGGC 422

Search completed: Sat Nov 14 01:13:28 1998
Job time : 2043 secs.

WQSEFH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 12:01:32 1998; MasPar time 11.52 Seconds
512.648 Million cell updates/sec

Tabular output not generated.

Title: >US-08-887-977-10
Description: (1-365) from US08887977.pap
Perfect Score: 2779
Sequence: 1 MFSTPVKIIILCQSIHLITQL.....NISQTSATADNDNASSFTM 365

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 3:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 34.572; Variance 151.305; scale 0.228

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	2779	100.0	365	29	W48086 Human dendritic cell	1.21e-255
2	990	35.6	355	15	R80950 Recombinant high affi	3.34e-81
3	990	35.6	355	6	R28272 Sequence in a high af	3.34e-81
4	984	35.4	350	12	R68811 Interleukin-8 recepto	1.26e-80
5	984	35.4	350	15	R80951 Recombinant high affi	1.26e-80
6	984	35.4	350	5	R80756 Interleukin 8 recepto	1.26e-80
7	983	35.4	350	5	R27791 Interleukin-8 recepto	1.57e-80
8	984	35.4	1060	14	R70123 IL8-R type 1-GPB 130	1.26e-80
9	965	34.7	358	15	R80952 Recombinant high affi	8.48e-79
10	961	34.6	358	11	R53745 Partial sequence of s	2.05e-78
11	961	34.6	410	11	R53743 Putative seven transm	2.05e-78
12	958	34.5	378	11	R53744 Putative seven transm	3.99e-78
13	951	34.2	355	7	R33420 Human IL-8 receptor f	1.88e-77
14	951	34.2	360	6	R28273 Sequence in a low affi	1.88e-77
15	951	34.2	360	15	R80758 Interleukin 8 recepto	1.88e-77
16	951	34.2	1064	14	R70124 IL8-R type 2-GPB 130	1.88e-77
17	948	34.1	359	11	R53747 Seven transmembrane r	3.65e-77
18	927	33.4	360	15	R80953 Recombinant high affi	3.79e-75

19	928	33.4	378	10	R54079 Epstein Barr virus in	3.04e-75
20	922	33.2	360	19	R99274 Chemokine receptor K5	1.55e-74
21	915	32.9	358	6	R28274 Sequence in a low af	5.38e-74
22	906	32.6	312	19	W02689 G-protein coupled hum	3.93e-73
23	906	32.6	312	16	R48717 G-protein coupled hum	3.93e-73
24	839	30.2	355	25	W26588 Human MIP-1 alpha/RAN	1.03e-66
25	839	30.2	355	11	R52749 C-C chemokine recepto	1.03e-66
26	839	30.2	355	24	W25751 Human MIP-1alpha/RAN	1.03e-66
27	836	30.1	354	11	R53932 Interleukin 8 (IL-8)	2.00e-66
28	812	29.2	415	23	W19780 Human G-protein chemo	3.95e-64
29	797	28.7	355	19	W03378 CC-chemokine receptor	1.07e-62
30	794	28.6	355	28	W31850 Human eosinophil eota	2.07e-62
31	794	28.6	355	19	W03377 CC-chemokine receptor	2.07e-62
32	794	28.6	355	24	W27124 Human chemokine recep	2.07e-62
33	793	28.5	355	19	W03376 CC-chemokine receptor	2.58e-62
34	793	28.5	355	23	W10100 Human C-C chemokine r	2.58e-62
35	790	28.4	352	24	W27125 Macaque chemokine rec	4.98e-62
36	780	28.1	356	26	W25943 Human CCR3 chemokine	4.48e-61
37	776	27.9	352	27	W27407 Human CCR5	1.08e-60
38	776	27.9	352	24	W27123 Human chemokine recep	1.08e-60
39	776	27.9	352	29	W23835 Human CC chemokine re	1.08e-60
40	775	27.9	355	24	W29179 Rat CC chemokine rece	1.35e-60
41	769	27.7	371	29	W23834 Human CC chemokine re	5.02e-60
42	764	27.5	352	20	W07602 Human G-protein chemo	1.51e-59
43	755	27.2	360	14	R79166 Human monocyte chemoa	1.09e-58
44	755	27.2	360	26	W35833 Human monocyte chemoa	1.09e-58
45	751	27.0	374	14	R79165 Human monocyte chemoa	2.61e-58

ALIGNMENTS

RESULT 1
ID W48086 standard; Protein; 365 AA.
AC W48086;
DT 11-JUN-1998 (first entry)
DE Human dendritic cell chemokine receptor.
KW Human; thymus expressed chemokine; TECK; MIP-3alpha; MIP-3beta;
OS Homo sapiens.
FH Key Location/Qualifiers
FT MISC_difference 193
FT /note= "encoded by CAN"
PN W09801557-A2.
PD 15-JAN-1998.
PF 02-JUL-1997; U10819.
PR 04-JUN-1997; US-048593.
PR 05-JUL-1996; US-675814.
PR 11-OCT-1996; US-028329.
PA (SCHE) SCHERING CORP.
PI Gish KC, Schall TJ, Vicari A, Wang W, Zlotnik A;
DR WPI; 98-101054/09.
DR N-PSDB; V15418.
PT Novel chemokines, e.g. thymus expressed chemokine - used for
PT treating inflammatory conditions including asthma.
PS Claim 3; Page 94-95; 202pp; English.
CC The present sequence represents human dendritic cell chemokine receptor.
CC Antibodies which bind to the protein can be used in detecting or
CC diagnosing various immunological conditions related to expression
CC of the protein. The nucleic acid can be used for screening and
CC isolating DNA clones for the chemokines, especially from other
CC species. The chemokine can be used in the treatment of conditions
CC associated with abnormal physiology or development, including
CC inflammatory conditions such as asthma.
SQ Sequence 365 AA;

Query Match 100.0%; Score 2779; DB 29; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.21e-255;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mfstpvkiiilcqsilhitqlilrcycapcrssgpgylrlyslcivlgilgnlivi 60
|||||
QY 1 MFSTPVKIIILCQSIHLITQLILRCYCAPCRSSGPGYLYRIAYSLICVLGILGNILWI 60
|||||

Db 61 tfafykkarsmtvdyllnmaiaidflvtlfpwvshatgawvfnscatckllkciyainf 120
 QY 61 TFAYFKKARSMTDVYLLNMAIADLFLVTLPFWAVSHATGAWVFNATCKLLKCIYAINF 120
 Db 121 ncgmllltcismdryiaivqatksfrlrsrtlprskliiclvvlgslvssstfvfnqky 180
 QY 121 NCGMLLTCISMDRYIAIVQATKSFRLSRTLPRSKLIICLVVGLSVIISSTFVFNQKY 180
 Db 181 ntqgsdvcepkytvspeirwklmgllellffgfiplmfimficytftvktlvqaqskr 240
 QY 181 NTQGSDVCEPKYTVSPEIRWKLMLGLELLFFGFIPLMFIMFICYTFTVKTlvQAQSKR 240
 Db 241 hkairvliavlvflacqiphnmvllvtaanlqkmrscsekligytktvtevlafhc 300
 QY 241 HKAIRVLIIVLVFLACQIPHNMVLLVTAANLQKMNRSCSEKLIGYTKTVTEVLAFHC 300
 Db 301 clnpvlyafiqgkfrnyflkildwcvrrkysgscagrysenisrqtsetadndna 360
 QY 301 CLNPVLYAFIQGKFRNYFLKILDWCVRRKYSGSCAGRYSENISRQTSETADNDNA 360
 Db 361 ssftm 365
 QY 361 SSFTM 365

RESULT 2

ID R80950 standard; Protein; 355 AA.
 AC R80950;
 DT 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotising vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Oryctolagus cuniculus.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PF 09-MAR-1995; U03032.
 PR 15-MAR-1994; US-210250.
 PR 02-MAY-1994; US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Iarosa GJ, Navarro J, Thomas KM;
 PI Witt DP;
 DR WPI; 95-336945/43.
 DR N-PSDB; Q99949.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 PT useful for treating inflammatory disorders, for detecting
 PT neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 2; Fig 1A-B; 74pp; English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is
 CC a high affinity receptor and the B subtype receptor (IL-8B) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 355 AA;

Query Match 35.6%; Score 990; DB 15; Length 355;
 Best Local Similarity 45.1%; Pred. No. 3.34e-81;
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;

Db 45 yvvviyalvflilgslvmlvlysrnsrvtvdyllnlnamadllfaltmpiwavsk 104
 QY 38 YLYRIAYSLICVLGLGNLWVITFAFYKARSMTDYLNNMAIADFLVTLTPFWAVSH 97
 Db 105 ekg-wifgtplckvslvkevnfygilllacisvdrlyaiavhatrltqk-rhl--vkf 160
 QY 98 ATGAWVFNATCKLLKGIYAINFCNCGMLLTCISMDRYIAIVQATKSFRLSRTLPRSKI 157

Db 161 iclgiwalslilsipfflgrqvfpnnssprvc---yedghntakwrmvrlilphtfgfi 217
 QY 158 ICLVWGLSVIISSTFVFNQKYNKTQGS-D-VCEPKYQIVS-EPIRWKLLMLGLELLFGFF 215
 Db 218 lpllvmlfcygtlrltfqahmgqkhramrvifavvllfllcwlpynlvlladtlmrthv 277
 QY 216 IPLMFIMFCYTFIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNNVLLVTAANLGM 275
 Db 278 iqetqrrndidraldateilgflhsclnpiiyafiqnfrngfklml 325
 QY 276 NR-SCQSEKLGITKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLKIL 322

RESULT 3

ID R28272 standard; Protein; 355 AA.
 AC R28272;
 DT 04-APR-1993 (first entry)
 DE Sequence in a high affinity recombinant rabbit interleukin-8
 DE (IL-8) receptor polypeptide in F3R.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Oryctolagus cuniculus.
 PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992; U02977.
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Navarro J, Thomas KM, Witt DP;
 DR WPI; 92-382123/46.
 DR N-PSDB; Q30011.
 PT Recombinant mammalian interleukin-8 receptor - used for screening
 PT interleukin-8 binding antagonists, used to treat inflammation
 PS Claim 2; Fig 1; 7lpp; English.
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments.
 SQ Sequence 355 AA;

Query Match 35.6%; Score 990; DB 6; Length 355;
 Best Local Similarity 45.1%; Pred. No. 3.34e-81;
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;

Db 45 yvvviyalvflilgslvmlvlysrnsrvtvdyllnlnamadllfaltmpiwavsk 104
 QY 38 YLYRIAYSLICVLGLGNLWVITFAFYKARSMTDYLNNMAIADFLVTLTPFWAVSH 97
 Db 105 ekg-wifgtplckvslvkevnfygilllacisvdrlyaiavhatrltqk-rhl--vkf 160
 QY 98 ATGAWVFNATCKLLKGIYAINFCNCGMLLTCISMDRYIAIVQATKSFRLSRTLPRSKI 157
 Db 161 iclgiwalslilsipfflgrqvfpnnssprvc---yedghntakwrmvrlilphtfgfi 217
 QY 158 ICLVWGLSVIISSTFVFNQKYNKTQGS-D-VCEPKYQIVS-EPIRWKLLMLGLELLFGFF 215
 Db 218 lpllvmlfcygtlrltfqahmgqkhramrvifavvllfllcwlpynlvlladtlmrthv 277
 QY 216 IPLMFIMFCYTFIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNNVLLVTAANLGM 275
 Db 278 iqetqrrndidraldateilgflhsclnpiiyafiqnfrngfklml 325
 QY 276 NR-SCQSEKLGITKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLKIL 322

DR N-PSDB; Q99006.
 PT New antibodies against interleukin 8 type B receptor - used to treat
 PT or prevent inflammation, also for detecting receptor expression and
 PT purification.
 PS Example 2; Columns 41-44; 62pp; English.
 CC Antibodies directed against the interleukin-8 receptor B can be used
 CC to treat or prevent inflammation e.g. psoriasis, dermatitis,
 CC rheumatoid arthritis and particularly inflammatory bowel disease and
 CC chronic lung inflammation. When immobilised, these antibodies may
 CC be used to detect interleukin-8 receptor B expression in cells and
 CC tissues and for affinity purification of interleukin-8 receptor B
 CC from cells. A sequence encoding amino acids 23-314 of this fragment
 CC of the high affinity interleukin-8 receptor, was used to probe lambda
 CC cgt10 cDNA libraries for the human interleukin-8 type B receptor.
 SQ Sequence 350 AA;

Query Match 35.4%; Score 984; DB 15; Length 350;
 Best Local Similarity 44.4%; Pred. No. 1.26e-80;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 y-spcmetetlnkyvviayalvlllsgnslvmlvlysvgrsvtdvylalnalaad 85
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNILVITFAFYKARSMYDVLNMAIAD 83
 Db 86 llfaltlpiwaaskvng-wifgtfclkvsvllkevnfygilllacisvdrlylaivhtr 144
 QY 84 ILFVLTPFWAVSHATGAWVFSNATCKLKGIIYAINFNCGMILLTCISMDRYIAIVQATK 143
 Db 145 tltdq-rhl--vkfvcgclgwmnslpfflfrqayhpnsspyc---yevlgndtakw 198
 QY 144 SFRLRSRTLPKSKIIICLVWGLSVIISSTFVFNQKYNQGS-D-VCEPKYQTVS-EPH 201
 Db 199 rmvlriltphgtfivplfvmfvcygtlrltkfahmgqkhrmvrfavvllfclwlp 258
 QY 202 KLLMLGLELLFGFFIPLMFMEFCTYFIVKTLVQAQNSKRHKRAIRVIAVVLFLACQIPH 261
 Db 259 nlvlladtlnrtqvigetcehnnigraldateilgflhscnlpiyafignfrhgf 318
 QY 262 NMVLLVTAANLGRMNR-SCOSEKLIGYTKTVTEVLAFHCLLPVLYAFIGQFRNYFLK 320
 Db 319 il 320
 QY 321 IL 322

RESULT 7
 ID R2791 standard; Protein; 350 AA.

AC R2791;
 DT 12-MAR-1993 (first entry)
 DE Interleukin-8 receptor.
 KW IL-8; G-protein coupled receptor family; rhodopsin superfamily;
 KW pro-inflammatory cytokine.
 OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	domain	40..67
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	76..98
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	112..133
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	155..174
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	204..226
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	243..264
FT	/label=	transmembrane
FT	/note=	"putative"
FT	domain	292..312

FT /label= transmembrane
 FT /note= "putative"

PN W09217497-A.
 PD 15-OCT-1992.
 PR 23-MAR-1991; U02317.
 PR 29-MAR-1991; US-677211.
 PR 19-DEC-1991; US-810782.
 PA (GETH) GENENTECH INC.
 PI Holmes WE, Lee J, Wood WI;
 WPI; 92-366191/44.
 DR N-PSDB; Q29505.
 PT Isolated human platelet factor 4 super-family receptor
 PT polypeptide and corresp. antibodies and DNA - useful as
 PT diagnostic and screening agents, and for treating inflammation or
 PT PF4AR-mediated disorders
 PS Claim 7; Fig 2; 78pp; English.
 CC A cDNA library constructed from human neutrophil mRNA in the mammalian
 CC expression vector PRK5B was transfected into COS-7 cells as pools of
 CC 2500 clones. One positive pool from the first 58 transfections was
 CC partitioned into smaller pools until a pure clone (PRK5B.118r1.1)
 CC was obtained. The ORF encodes a protein of 350 amino acids which
 CC shares several features with the G-protein coupled receptors of the
 CC rhodopsin superfamily, including 7 hydrophobic (transmembrane)
 CC domains. The IL-8 receptor is a preferred PF4AR superfamily member
 CC of the invention. See also Q29506 and Q37107.
 SQ Sequence 350 AA;

Query Match 35.4%; Score 983; DB 5; Length 350;
 Best Local Similarity 44.4%; Pred. No. 1.57e-80;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 y-spcmetetlnkyvviayalvlllsgnslvmlvlysvgrsvtdvylalnalaad 85
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNILVITFAFYKARSMYDVLNMAIAD 83
 Db 86 llfaltlpiwaaskvng-wifgtfclkvsvllkevnfygilllacisvdrlylaivhtr 144
 QY 84 ILFVLTPFWAVSHATGAWVFSNATCKLKGIIYAINFNCGMILLTCISMDRYIAIVQATK 143
 Db 145 tltdq-rhl--vkfvcgclgwmnslpfflfrqayhpnsspyc---yevlgndtakw 198
 QY 144 SFRLRSRTLPKSKIIICLVWGLSVIISSTFVFNQKYNQGS-D-VCEPKYQTVS-EPH 201
 Db 199 rmvlriltphgtfivplfvmfvcygtlrltkfahmgqkhrmvrfavvllfclwlp 258
 QY 202 KLLMLGLELLFGFFIPLMFMEFCTYFIVKTLVQAQNSKRHKRAIRVIAVVLFLACQIPH 261
 Db 259 nlvlladtlnrtqvigetcehnnigraldateilgflhscnlpiyafignfrhgf 318
 QY 262 NMVLLVTAANLGRMNR-SCOSEKLIGYTKTVTEVLAFHCLLPVLYAFIGQFRNYFLK 320
 Db 319 il 320
 QY 321 IL 322

RESULT 8

ID R70123 standard; Protein; 1060 AA.
 AC R70123;
 DT 14-FEB-1996 (first entry)
 DE IL8-R type 1-GPB 130 fusion protein.
 KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;
 KW red blood cell; cytokine receptor; glycoprotein binding peptide 130;
 KW GBP 130; GBP; glycoprotein binding peptide homologue; glycoprotein A.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 PN W09506737-A.
 PD 09-MAR-1995.
 PF 01-SEP-1994; G01900.
 PR 03-SEP-1993; GB-018350.
 PR 23-AUG-1994; GB-017021.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;

[illegible]

OS Homo sapiens.
 PN WO9306229-A.
 PD 01-APR-1993.
 PF 14-SEP-1992; U07641.
 PR 13-SEP-1991; US-759568.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Murphy PM;
 DR WPI: 93-117549/14.
 DR N-PSDB; Q38747.
 PT New interleukin-8 receptor aminoacid sequence - and corresp. cDNA
 expressed in Xenopus laevis oocytes or transfected host cells,
 PT for screening ligands of IL-8 receptor and gene therapy
 PS Claim 1; Fig 3; 39pp; English.
 CC cDNA libraries from 2 and 3.5 kb fractions of poly(A)+ RNA from HL60
 CC neutrophils sep'd. of a sucrose gradient were made in UniZAP. The
 CC libraries were screened with F3R oligonucleotide probe (from rabbit
 CC IL-8 receptor) and under low stringency with a p2 cDNA probe
 CC synthesised from random primers, to isolated the clone p2, encoding
 CC human IL-8 receptor protein. The presence or absence of the DNA
 CC coding IL-8R or related MIP-2 receptor may be detected using
 CC portions of the p2 clone as probes. P2 may also be used to screen
 CC for ligands of IL-8R and may also be used in gene therapy to treat a
 CC patient deficient in IL-8R. The IL-8R is a gro receptor and has
 CC intracellular calcium-mobilising and ligand-binding properties.
 SQ Sequence 355 AA;

Query Match 34.2%; Score 951; DB 7; Length 355;
 Best Local Similarity 43.8%; Pred. No. 1.88e-77;
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;
 Db 44 yfvviyalvllslqnslymvllyrvgrsvtdvllnlaladlalfaltlpiwaask 103
 QY 38 YLXRIAYSLICVLGLGNILVITFAFYKARSMTDVLNMAIADILFVLTLPFWAVSH 97
 Db 104 vng-wifgtfclckvsvllkevnfygillacisvdrylaivhatrltqk-ryl--vkf 159
 QY 98 ATGAWFSNATCKLLKGIYAINFNGMLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157
 Db 160 iclsiwgslslalpvllfrtvyssnvspac---yedmgnntanwmlrlilpqsgf 216
 QY 158 ICLVWGLSVIISSTFVFNQK-YNTQSDVCEPKYQVSE-PIRWKLLMLGLELFGFF 215
 Db 217 vpllmifcygftlrlfkahmgqkhravrfavvllfllcwlpylnvlladtlmrtqv 276
 QY 216 IPLMFIMFCYTFIVKTLVQAQSKRHKRAIRVIAVFLACQIPHNVLVTAANLGKM 275
 Db 277 iqetcerhndaldateilgilhscnlpiyafigqkfrhglkil 324
 QY 276 NR-SCQSEKLGTYKTVTEVLAFLHCLLPVLYAFIGQKFRNYFLKIL 322

RESULT 14
 ID R28273 standard; Protein; 360 AA.
 AC R28273;
 DT 04-APR-1993 (first entry)
 DE Sequence in a low affinity recombinant human interleukin-8
 DE (IL-8) receptor polypeptide in 4AB.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Homo sapiens.
 PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992;
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPK) REPLIGEN CORP.
 PA (DYBO-) UNIV BOSTON.
 PI Navarro J, Thomas KM, Witt DP;
 DR N-PSDB; Q30012.
 PT Recombinant mammalian interleukin-8 receptor - used for screening
 PT interleukin-8 binding antagonists, used to treat inflammation
 PS Disclosure; Fig 2; 71pp; English.

CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA. to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments. A human
 CC peripheral blood leukocyte lambda gt11 cDNA library (5' stretch) was
 CC screened with a 652 bp EcoRI/BamHI fragment (including nucleotides
 CC -27 to 825) of the rabbit F3R clone. After tertiary screening
 CC several human clones which hybridized to the rabbit IL-8 probe
 CC were isolated. The insert of one such clone, termed 4AB was
 CC sequenced (Q30012).
 SQ Sequence 360 AA;

Query Match 34.2%; Score 951; DB 6; Length 360;
 Best Local Similarity 43.8%; Pred. No. 1.88e-77;
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;
 Db 49 yfvviyalvllslqnslymvllyrvgrsvtdvllnlaladlalfaltlpiwaask 108
 QY 38 YLXRIAYSLICVLGLGNILVITFAFYKARSMTDVLNMAIADILFVLTLPFWAVSH 97
 Db 109 vng-wifgtfclckvsvllkevnfygillacisvdrylaivhatrltqk-ryl--vkf 164
 QY 98 ATGAWFSNATCKLLKGIYAINFNGMLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157
 Db 165 iclsiwgslslalpvllfrtvyssnvspac---yedmgnntanwmlrlilpqsgf 221
 QY 158 ICLVWGLSVIISSTFVFNQK-YNTQSDVCEPKYQVSE-PIRWKLLMLGLELFGFF 215
 Db 222 vpllmifcygftlrlfkahmgqkhravrfavvllfllcwlpylnvlladtlmrtqv 281
 QY 216 IPLMFIMFCYTFIVKTLVQAQSKRHKRAIRVIAVFLACQIPHNVLVTAANLGKM 275
 Db 282 iqetcerhndaldateilgilhscnlpiyafigqkfrhglkil 329
 QY 276 NR-SCQSEKLGTYKTVTEVLAFLHCLLPVLYAFIGQKFRNYFLKIL 322

RESULT 15
 ID R80758 standard; Protein; 360 AA.
 AC R80758;
 DT 26-MAR-1996 (first entry)
 DE Interleukin 8 receptor B.
 KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;
 KW rheumatoid arthritis; inflammatory bowel disease;
 KW chronic lung inflammation; treatment; antibody;
 KW affinity purification; detection.
 OS Homo sapiens.
 PN US5440021-A.
 PD 08-AUG-1995.
 PF 29-MAR-1991; 677211.
 PR 29-MAR-1991; US-677211.
 PR 25-FEB-1994; US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HERBERT C.
 PA (KIMK/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J;
 DR WPI: 95-283151/37.
 DR N-PSDB; Q99008.
 PT New antibodies against interleukin 8 type B receptor - used to treat
 PT or prevent inflammation, also for detecting receptor expression and
 PT purification.
 PS Disclosure; Columns 53-56; 62pp; English.
 CC Antibodies directed against the interleukin-8 receptor B can be used
 CC to treat or prevent inflammation e.g. psoriasis, dermatitis,

SQ Sequence 360 AA;

Search completed: Fri Nov 13 12:03:22 1998
Job time : 110 secs.

W O R L D
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Nov 13 12:03:41 1998; MasPar time 16.03 Seconds
Tabular output not generated. 831.784 Million cell updates/sec

Title: >US-08-887-977-10
Description: (1-365) from US08887977.ppe
Perfect Score: 2779
Sequence: 1 MFSTPVKIIICQSLHTQL.....NISQTSETADNDNASSFTM 365

Scoring table: PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d
Statistics: Mean 47.004; Variance 114.799; scale 0.409

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2445	88.0	369	2	JC5068	G protein-coupled rec
2	990	35.6	355	2	JQ1231	interleukin-8 recepto
3	986	35.5	350	2	A39445	interleukin-8 recepto
4	965	34.7	358	2	A53752	interleukin-8 recepto
5	961	34.6	378	2	B55735	lymphocyte-specific G
6	959	34.5	378	2	A53735	G protein-coupled rec
7	951	34.2	360	2	A53611	interleukin-8 recepto
8	927	33.4	360	2	A57160	chemokine (C-C) recep
9	928	33.4	378	2	A45680	G protein-coupled pep
10	906	32.6	356	2	S42096	interleukin-8 recepto
11	891	32.1	359	2	A48921	interleukin-8 recepto
12	870	31.3	360	2	JC4587	chemokine (C-C) recep
13	839	30.2	355	2	A45177	chemokine (C-C) recep
14	836	30.1	354	2	A23669	interleukin-8 recepto
15	835	30.0	359	2	I49341	MIP-1 alpha receptor
16	793	28.5	355	2	G02436	C-C chemokine recepto
17	784	28.2	355	2	A57337	chemokine (C-C) recep
18	776	27.9	352	2	A43113	chemokine (C-C) recep
19	769	27.7	352	2	G02653	CC chemokine receptor
20	762	27.4	355	2	I49339	macrophage inflammato
21	755	27.2	360	2	JC2443	chemokine (C-C) recep
22	751	27.0	374	2	I38450	chemokine (C-C) recep
23	738	26.6	355	2	JC5067	G protein-coupled rec

24	734	26.4	354	2	I58186	hypothetical G-protein
25	726	26.1	374	2	S32785	G protein-coupled rec
26	723	26.0	359	2	S15403	angiotensin II recept
27	721	25.9	355	2	JC4304	orphan G-protein-coup
28	717	25.8	352	2	A45747	leukocyte-derived sev
29	716	25.8	359	2	A42656	angiotensin II recept
30	713	25.7	359	2	JH0621	angiotensin receptor
31	715	25.7	374	2	S42628	G-protein coupled rec
32	711	25.6	359	2	JC1104	angiotensin II recept
33	708	25.5	352	2	G00048	angiotensin II recept
34	709	25.5	359	2	JQ1516	angiotensin II recept
35	706	25.4	353	2	S28787	neuropeptide Y/peptid
36	707	25.4	356	2	I49340	MIP-1 alpha receptor
37	707	25.4	359	2	JC2134	angiotensin II recept
38	699	25.2	359	2	JC1193	angiotensin II recept
39	698	25.1	350	2	JN0621	G-protein coupled rec
40	698	25.1	359	2	S44425	angiotensin II recept
41	697	25.1	359	2	A48857	AT1 angiotensin II re
42	690	24.8	359	2	JC1194	angiotensin II recept
43	687	24.7	327	2	S56162	MCDR15 protein - huma
44	687	24.7	372	2	S26667	G protein-coupled rec
45	683	24.6	383	2	S55594	G protein-coupled rec

ALIGNMENTS

RESULT 1
ENTRY JC5068 #type complete
TITLE G protein-coupled receptor CKR-L3 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 10-Sep-1997
ACCESSIONS JC5068
REFERENCE JC5067
#authors Zaballios, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
#journal Biochem. Biophys. Res. Commun. (1996) 227:846-853
#title Molecular cloning and RNA expression of two new human chemokine receptor-like genes.
#accession JC5068
#molecule_type DNA
#residues 1-369 #label ZAB
#cross-references EMBL:Z79784; NID:gl668737; PID:e264774; PID:gl668738
COMMENT This protein belongs to the family of alpha chemokine receptors.
KEYWORDS G protein-coupled receptor; transmembrane protein
FEATURE
42-68 #domain transmembrane #status predicted #label TM1\
79-99 #domain transmembrane #status predicted #label TM2\
115-136 #domain transmembrane #status predicted #label TM3\
160-180 #domain transmembrane #status predicted #label TM4\
212-233 #domain transmembrane #status predicted #label TM5\
250-271 #domain transmembrane #status predicted #label TM6\
292-315 #domain transmembrane #status predicted #label TM7
SUMMARY
#length 369 #molecular-weight 42002 #checksum 1020

Query Match 88.0%; Score 2445; DB 2; Length 369;
Best Local Similarity 98.8%; Pred. NO. 0.00e+00;
Matches 324; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
Db 42 LFVPDIATSLICVFGLLGNILWITFAFYKKARSMTDVYLLNMAIADILFVLTPFWAVSH 101
QY 39 LY-RIATSLICVLGLLGNILWITFAFYKKARSMTDVYLLNMAIADILFVLTPFWAVSH 97
Db 102 ATGAWFVSNAATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSPRLSRTPRSKI 161
QY 98 ATGAWFVSNAATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSPRLSRTPRSKI 157
Db 162 ICLVWGLSVLISSSTFVFNQKYNQTCSDVCEPKYQVSEPIRWKLLMLGLLELFGFFIP 221
QY 158 ICLVWGLSVLISSSTFVFNQKYNQTCSDVCEPKYQVSEPIRWKLLMLGLLELFGFFIP 217
Db 222 LMFIFCYTIVKTLVQAKNSKRHKAIRVIAVVLVFLACQIPHNWLLVTAANLGMNR 281


```

Db 119 AK-SWIFGVYLCRGIGIYKLSFSGMLLLCISIDRYVAIVQAVSRHRHARVLLISKL 177
QY 98 ATGAWFSNATCKLKGIIYAINFCGMLLTCTISDMRYIAIVQATKSRFLRSRTLPKSKI 157

Db 178 SCVGIWMLALFLSIPPELLYSGLOKNGEDT--LRCSLVSQAVE-ALITIOVAQMVFGFLV 234
QY 158 ICLVWGLSVLISSSVFVFNQKNTQSDVCEPKYQVSPIRKWLMLGL-ELLFGFFI 216

Db 235 PMLAMSFCLIIIRITLLQARNFERNKAIRVIIAVVVVFIYFQLPYNGVLAQTVANFNIT 294
QY 217 PLMFMEFCYTFIVKTLVQAQNSKRHKRAIRVIIAVVFLACQIPHNMLLV-TAANLGKM 275

Db 295 NSCETSKOLNIADYDTYSLASVRCVNPFLYAFIGVFKFSDLFKFLKGLGCL 347
QY 276 NRSCQSEKLIGYTKTVEVLAFLHCLNPLVLYAFIGKFRNYFLKILKDLWCV 328

RESULT 7
ENTRY A53611 #type complete
TITLE Interleukin-8 receptor type B - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change
29-Aug-1997
ACCESSIONS 137898; 138712; A53611; A39446
REFERENCE 137898
#authors Anuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
#journal J. Biol. Chem. (1994) 269:26381-26389
#title Comparison of the genomic organization and promoter function
for human interleukin-8 receptors A and B.
#cross-references MUID:95014476
#accession 137898
#status preliminary
#molecule_type DNA
#residues 1-360 #label RES
#cross-references EMBL:U11869; NID:G511801; PID:G511803
#accession 138712
#status preliminary
#molecule_type RNA
#residues 1-15 #label RE2
#cross-references EMBL:U11872; NID:G511808; PID:G511809; EMBL:U11873;
NID:G511810; PID:G511811; EMBL:U11874; NID:G511812;
PID:G511813; EMBL:U11875; NID:G511814; PID:G511815;
EMBL:U11876; NID:G511816; PID:G511817; EMBL:U11877;
NID:G511818; PID:G511819; EMBL:U11878; NID:G511820;
PID:G511821
#accession A53611
#authors Sprenger, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.I.;
Kelvin, D.J.
#journal J. Biol. Chem. (1994) 269:11065-11072
#title Structure, genomic organization, and expression of the human
interleukin-8 receptor B gene.
#accession A53611
#status preliminary
#molecule_type DNA
#residues 6-360 #label SPR
#cross-references GB:M99412; GB:L19593
#accession A39446
#authors Murphy, P.M.; Tiffany, H.L.
#journal Science (1991) 253:1280-1283
#title Cloning of complementary DNA encoding a functional human
interleukin-8 receptor.
#cross-references MUID:91368200
#accession A39446
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 6-360 #label MUR
#cross-references GB:M73969
COMMENT This receptor, unlike IL8RA, binds several peptides besides
interleukin-8, including GRO, NAP-2, and ENA-78.
GENETICS
#gene GDB:IL8RB; IL8RA
#cross-references GDB:I27868; OMIM:146928

```

```

#map_position 2q35-2q35
KEYWORDS G protein-coupled receptor; transmembrane protein
SUMMARY #length 360 #molecular_weight 40759 #checksum 3062

Query Match 34.2%; Score 951; DB 2; Length 360;
Best Local Similarity 43.8%; Pred. No. 138e-133;
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;

Db 49 YFVVIYALVFLSLGSLNMLVILYISRGVSRVTDVYLLNLALADLFLALFLPIWAASK 108
QY 38 YLYRIAYSGLICVLGSLNGLVITFAFYKARSMDTVYLLNMAIDILFVLTPFWAVSH 97

Db 109 VNG-WIEFTFFCKVSVLLKEVNFYSGILLACISVDRIYAIYHATRTLTOK-RYL--VKF 164
QY 98 ATGAWFSNATCKLKGIIYAINFCGMLLTCTISDMRYIAIVQATKSRFLRSRTLPKSKI 157

Db 165 ICLSIWGLSLLLALPVLFLFRRTVYSSNVSPAC--YEDMGNNNTANRWMLRLRLPQSGFI 221
QY 158 ICLVWGLSVLISSSVFVFNQK-YNTQSDVCEPKYQVSE-PIRWKLLMLGLELLFGFF 215

Db 222 VPLLMLFCYGTTLRTLFKAHMGQKRAMRVIFAVVFLFLCWLPLNVLVLLADTLMTQV 281
QY 216 IPLMFMEFCYTFIVKTLVQAQNSKRHKRAIRVIIAVVFLACQIPHNMLLVTAANLGKM 275

Db 282 IOETCERNHHDRAIDALDAEILGILHSLNPLIYAFIGKFRHGLKIL 329
QY 276 NR-SCQSEKLIGYTKTVEVLAFLHCLNPLVLYAFIGKFRNYFLKIL 322

RESULT 8
ENTRY A57160 #type complete
TITLE Chemokine (C-C) receptor 4 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change
15-Aug-1997
ACCESSIONS A57160
REFERENCE A57160
#authors Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf,
A.J.; Proudfoot, A.E.I.; Wells, T.N.C.
#journal J. Biol. Chem. (1995) 270:19495-19500
#title Molecular cloning and functional expression of a novel CC
chemokine receptor cDNA from a human basophilic cell line.
#accession A57160
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1-360 #label POW
#cross-references GB:X85740
#note source clone K5-5
GENETICS
#gene GDB:CMKBR4
#cross-references GDB:677463
#map_position 3p21-3p21
KEYWORDS G protein-coupled receptor; glycoprotein; phosphoprotein;
transmembrane protein

FEATURE
40-65 #domain transmembrane #status predicted #label TM1\
76-97 #domain transmembrane #status predicted #label TM2\
112-133 #domain transmembrane #status predicted #label TM3\
151-175 #domain transmembrane #status predicted #label TM4\
208-226 #domain transmembrane #status predicted #label TM5\
243-264 #domain transmembrane #status predicted #label TM6\
291-308 #domain transmembrane #status predicted #label TM7\
29-276,110-187 #disulfide_bonds #status predicted\
72,350 #binding_site phosphate (Ser) (covalent) (by casein
kinase II) #status predicted\
145 #binding_site phosphate (Ser) (covalent) (by protein
kinase C) #status predicted\
183,194 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
321 #binding_site phosphate (Thr) (covalent) (by protein
kinase C) #status predicted\
SUMMARY #length 360 #molecular_weight 41402 #checksum 4713

```



```
##status      preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues    1-359 #label RES
##cross-references EMBL:028406; NID:g881551; PID:g881552
SUMMARY      #length 359 #molecular-weight 41857 #checksum 8333

Query Match      30.0%; Score 835; DB 2; Length 359;
Best Local Similarity 38.4%; Pred.No. 5.47e-114;
Matches 108; Conservative 89; Mismatches 76; Indels 8; Gaps 7;

Db 45 YSLVFIIGLGNMVLILIKYRKQLQIMTNIYLFNLAISDILFLFTVPFW-IHYVLWNEW 103
   |||: ::|||::|||: : ||: : |||::||: ||: ||: ||: ||: ||: |
Qy 44 YSLICVLGLGNILVITFAFYKKARSMTDVYLLNMAIADILFVLTLPFWAVSHAT-GAW 102
   ||: ||: ||: ||: : : :: : ||: ||: ||: ||: ||: ||: ||: :

Db 104 GFCHYCKMLSGFYLYALYSEIFFIILLTIDRYLAIVHAV--FALRARTVTFATITSIT 161
   ||: ||: ||: ||: : : :: : ||: ||: ||: ||: ||: ||: ||: :
Qy 103 VFSNATCKLLKGIYAINFCNGMLLLTCISMDRYAIQVATKSPRLRSKTLPRSKIICLV 162
   |||: :: : ||: :: : ||: ||: ||: ||: ||: ||: ||: ||: ||: |

Db 162 WGLAGLAALPEEFIFHESQDSFGESCSPRYPE-GEEDSWKRFHARMNI-FGLALPLLYM 219
   |||: :: : ||: :: : ||: ||: ||: ||: ||: ||: ||: ||: ||: |
Qy 163 WGLSVIISSSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWK-LLMGLLELLFGFFIPLFM 221
   |||: :: : ||: :: : ||: ||: ||: ||: ||: ||: ||: ||: ||: |

Db 220 VICYSGIITLLRCPNKKKKAIRLIFWVMIVFFIFWTYPYNLVLLFSAFHSTFLETSC 279
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 222 IFCTYFIVKTLVQAQSKRKAIRVIAVVLVFLACOIPHNWVLLVTAANLGKWRSCQS 281
   | : : |||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

Db 280 SKHLDLAMOQTEVIAYTHCCVNPVIYAFVGERFRKH-LRLF 319
   | : : |||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 282 EKLIGYTKTVTEVLAFHCCCLNPVLYAFIGKPRNYFLKIL 322
```

Search completed: Fri Nov 13 12:04:53 1998
Job time : 72 secs.

W P S R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 12:05:10 1998; MasPar time 10.77 Seconds
Tabular output not generated. 850.044 Million cell updates/sec

Title: >US-08-887-977-10
Description: (1-365) from US08887977.ppt
Perfect Score: 2779
Sequence: 1 MFSTPVKILLCQSILHTQL.....NISROTSETADNDNASSFTM 365

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 49.069; Variance 98.526; scale 0.498

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2445	88.0	369	1	CCR6_HUMAN C-C CHEMOKINE RECEPTOR	0.00e+00
2	994	35.8	350	1	IL8A_PANTR HIGH AFFINITY INTERLEU	1.02e-170
3	994	35.8	350	1	IL8A_GORGO HIGH AFFINITY INTERLEU	1.02e-170
4	994	35.8	350	1	IL8A_PORPY HIGH AFFINITY INTERLEU	1.02e-170
5	990	35.6	355	1	IL8A_RABIT HIGH AFFINITY INTERLEU	6.81e-167
6	984	35.4	350	1	IL8B_HUMAN HIGH AFFINITY INTERLEU	1.17e-168
7	978	35.2	360	1	IL8B_BOVIN HIGH AFFINITY INTERLEU	2.02e-167
8	965	34.7	358	1	IL8B_RABIT HIGH AFFINITY INTERLEU	9.57e-165
9	961	34.6	378	1	CCR7_HUMAN C-C CHEMOKINE RECEPTOR	6.36e-164
10	959	34.5	353	1	IL8B_MACMU HIGH AFFINITY INTERLEU	1.64e-163
11	959	34.5	378	1	CCR7_MOUSE C-C CHEMOKINE RECEPTOR	1.64e-163
12	951	34.2	353	1	IL8B_PANTR HIGH AFFINITY INTERLEU	7.25e-162
13	951	34.2	360	1	IL8B_HUMAN HIGH AFFINITY INTERLEU	7.25e-162
14	945	34.0	353	1	IL8B_GORGO HIGH AFFINITY INTERLEU	1.24e-160
15	927	33.4	360	1	CCR4_HUMAN C-C CHEMOKINE RECEPTOR	6.19e-157
16	920	33.1	359	1	IL8B_RAT HIGH AFFINITY INTERLEU	1.69e-155
17	911	32.8	349	1	IL8A_RAT HIGH AFFINITY INTERLEU	1.19e-153
18	894	32.2	357	1	CCR6_HUMAN PROBABLE G PROTEIN-COU	3.64e-150
19	891	32.1	359	1	IL8B_MOUSE HIGH AFFINITY INTERLEU	1.50e-149
20	870	31.3	360	1	CCR4_MOUSE C-C CHEMOKINE RECEPTOR	2.99e-145
21	839	30.2	355	1	CCR1_HUMAN C-C CHEMOKINE RECEPTOR	6.49e-139
22	835	30.0	359	1	CCR3_MOUSE PROBABLE C-C CHEMOKINE	4.26e-138
23	824	29.7	368	1	CCR3_HUMAN C-X-C CHEMOKINE RECEPT	7.48e-136

24	797	28.7	354	1	CCR5_RAT	C-C CHEMOKINE RECEPTOR	2.37e-130
25	794	28.6	355	1	CCR3_HUMAN	C-C CHEMOKINE RECEPTOR	9.68e-130
26	792	28.5	352	1	CCR5_MACMU	C-C CHEMOKINE RECEPTOR	2.47e-129
27	776	27.9	352	1	CCR5_HUMAN	C-C CHEMOKINE RECEPTOR	4.43e-126
28	770	27.7	373	1	CCR2_MOUSE	C-C CHEMOKINE RECEPTOR	7.33e-125
29	759	27.3	354	1	CCR5_MOUSE	C-C CHEMOKINE RECEPTOR	1.25e-122
30	757	27.2	355	1	CCR1_MOUSE	C-C CHEMOKINE RECEPTOR	3.19e-122
31	751	27.0	374	1	CCR2_HUMAN	C-C CHEMOKINE RECEPTOR	5.86e-121
32	738	26.6	355	1	GCY6_HUMAN	PROBABLE G PROTEIN-COU	2.27e-118
33	734	26.4	354	1	GPDR_RAT	PROBABLE G PROTEIN-COU	1.46e-117
34	726	26.1	374	1	BLR1_RAT	BURKITT'S LYMPHOMA REC	6.09e-116
35	723	26.0	359	1	AG2R_BOVIN	TYPE-1 ANGIOTENSIN II	2.46e-115
36	721	25.9	355	1	GPDR_HUMAN	PROBABLE G PROTEIN-COU	6.36e-115
37	717	25.8	352	1	CCR4_HUMAN	SDF-1 RECEPTOR (STROMA	4.03e-114
38	716	25.8	359	1	AG2S_RAT	TYPE-1B ANGIOTENSIN II	6.41e-114
39	713	25.7	359	1	AG2R_MOUSE	TYPE-1A ANGIOTENSIN II	2.59e-113
40	715	25.7	374	1	BLR1_MOUSE	BURKITT'S LYMPHOMA REC	1.02e-113
41	711	25.6	359	1	AG2R_HUMAN	TYPE-1A ANGIOTENSIN II	6.57e-113
42	711	25.6	359	1	AG2R_PIG	TYPE-1 ANGIOTENSIN II	6.57e-113
43	706	25.4	353	1	CCR4_BOVIN	SDF-1 RECEPTOR (STROMA	6.72e-112
44	707	25.4	356	1	CCR4_MOUSE	C-C CHEMOKINE RECEPTOR	4.22e-112
45	707	25.4	359	1	AG2R_RAT	TYPE-1A ANGIOTENSIN II	4.22e-112

ALIGNMENTS

RESULT 1	CCR6_HUMAN	STANDARD;	PRG:	369 AA.
ID	P51684; Q92846; P78553;			
AC	01-OCT-1996 (REL. 34, CREATED)			
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	C-C CHEMOKINE RECEPTOR TYPE 6 (C-C CCR-6) (CCR6) (GPR-CY4) (CCR-L3) (CCR6)			
GN	CMKBR6 OR STRL22 OR GPR29.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	LAUTENS L.L., MODI W., BONNER T.I.;			
RL	SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97040707.			
RA	ZABALLOS A., VARONA R., GUTIERREZ J., LIND P., MARQUEZ G.;			
RL	BIOCHEM. BIOPHYS. RES. COMMUN. 227:846-853(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MCCOY R., PERLMUTTER D.H.;			
RL	SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	LIAO F., LEE H.H., FARBER J.M.;			
RL	GENOMICS 40:175-180(1997).			
CC	- FUNCTION: ORPHAN RECEPTOR.			
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	STRONG, TO IL-8 RECEPTORS.			
EMBL	U45984; G1245059; -			
EMBL	279784; E264774; -			
DR	EMBL; U60000; G1515435; ALT_INIT.			
DR	EMBL; U68030; G1870666; ALT_INIT.			
DR	EMBL; U68032; G1870669; ALT_INIT.			
DR	MM; 601835; -			
DR	PROSITE; PS00237; G-PROTEIN RECEPTOR; 1.			
KW	G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.			
FT	DOMAIN 1 42			
FT	TRANSMEM 43 69			
FT	DOMAIN 70 78			
FT	TRANSMEM 79 99			
FT	DOMAIN 100 114			

AC P21109;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (CXCR-1).
 GN IL8RA OR CXCR1.
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; LAGOMORPHA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91378994.
 RA BECKMANN M.P., MUNGER W.E., KOZLOSKY C., VANDENBOS T., PRICE V.,
 RA LYMAN S., GERARD N.P., GERARD C., CERRETTI D.P.,
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 179:784-789(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ALBINO; TISSUE=NEUTROPHILS;
 RX MEDLINE; 91056034.
 RA THOMAS K.M., PYUN H.Y., NAVARRO J.,
 RL J. BIOL. CHEM. 265:20061-20064 (1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=NEUTROPHILS;
 RX MEDLINE; 92148149.
 RA LEE J., KUANG W.-J., RICE G.C., WOOD W.I.;
 RL J. IMMUNOL. 148:1261-1264 (1992).
 CC -!- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: NEUTROPHILS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE RECEPTOR FOR
 CC FMET-LEU-PHE (N-FORMYL PEPTIDE RECEPTOR).
 DR EMBL; M74240; G165439; -;
 DR EMBL; M58021; G165443; -;
 DR EMBL; M82873; G165441; -;
 DR PIR; A23669; A23669.
 DR PIR; A46483; A46483.
 DR PIR; JQ1231; JQ1231.
 DR GCRDB; GCR_0107; -;
 DR GCRDB; GCR_0108; -;
 DR GCRDB; GCR_0298; -;
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 67 1 (POTENTIAL).
 FT DOMAIN 68 73 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 74 92 2 (POTENTIAL).
 FT DOMAIN 93 114 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 115 138 3 (POTENTIAL).
 FT DOMAIN 139 159 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 160 184 4 (POTENTIAL).
 FT DOMAIN 185 204 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 205 232 5 (POTENTIAL).
 FT DOMAIN 233 247 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 248 270 6 (POTENTIAL).
 FT DOMAIN 271 290 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 291 313 7 (POTENTIAL).
 FT DOMAIN 314 355 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 21 21 POTENTIAL.
 FT DISULFID 115 192 BY SIMILARITY.
 FT CONFLICT 90 111 DLFLFALPFWAVSKRGWIFG ->
 FT CONFLICT 146 147 PAFCPDHAYLRGKQKRLDER (IN REF. 2).
 FT CONFLICT 204 204 HA -> QS (IN REF. 2).
 FT CONFLICT 287 288 R -> C (IN REF. 2).
 FT CONFLICT 288 288 DI -> EL (IN REF. 2).
 FT CONFLICT 355 355 AA; 40622 MW; 2EB3947D CRC32;
 SQ SEQUENCE

Query Match 35.68; Score 990; DB 1; Length 355;
 Best Local Similarity 45.18; Pred. No. 6.81e-170;
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;
 DB 45 YVWVVIYALVFLSLGNSLVMLVILYSRSNRSVTDVYLLNLAMADLLFALTPIMAVSK 104
 QY 38 YLRIAYSGLCVGLGNLNVITFAFYKARSMTDVLNMAIDILFVLTLPFWAVSH 97
 DB 105 EKG-WIFGTPCKVSVLKVREVFYSGILLACISVDRLAIVHATRLTQK-RHL--VKF 160
 QY 98 ATGAWFVSATCKLLKGIYAINFCNCGMLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157
 DB 161 ICLGIWALSILSLPFLPQVFPSPNNSSPVC---YEDLGHNTAKRWVRILPHTGFI 217
 QY 158 ICLVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EPTRWKLMLGLLELFGFF 215
 DB 218 LPLVLMFCYGETLRTLFQAHMGQKHRAVEIVAVLVILLCWLPYNLVLLADPLMETHV 277
 QY 216 IPLMFMIFCYTFIVKTLVQONSKRAIRVIAVVLVFLACQLPHNNVLLVTAAANLGM 275
 DB 278 IQETCQRNRIDRALDATEILGELHSCNPIIYAFIQGNFNGFLKML 325
 QY 276 NR-SCOSEKLIQYTKVTVEVLAFHCLNPLVYAFIQGKERNYFLKIL 322
 RESULT 6
 ID IL8A_HUMAN STANDARD; PRT; 350 AA.
 AC P25024;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (IL-8 RECEPTOR TYPE
 DE 1) (CXCR-1) (CDW128).
 GN IL8RA OR CXCR1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91368199.
 RA HOLMES W.E., LEE J., KUANG W.-J., RICE G.C., WOOD W.I.;
 RL SCIENCE 253:1278-1280(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE; 93205012.
 RA CERRETTI D.P., KOZLOSKY C.J., VANDEN BOS T., NELSON N., GEARING D.P.,
 RA BECKMANN M.P.;
 RL MOL. IMMUNOL. 30:359-367(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93252387.
 RA MOLLEREAU C., PASSAGE E., MATTEI M.-G., VASSART G., PARMENTIER M.;
 RL GENOMICS 16:248-251(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE; 95014476.
 RA AHUTA S.K., SHETTY A., TIFFANY H.L., MURPHY P.M.;
 RL J. BIOL. CHEM. 269:26381-26389(1994).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE; 92355587.
 RA LEE J., HORUK R., RICE G.C., BENNETT G.L., CAMERATO T., WOOD W.I.;
 RL J. BIOL. CHEM. 267:16283-16287(1992).
 CC -!- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO MGSA (GRO) WITH A LOW AFFINITY.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

QY 98 ATGAWFVNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSFRLRSRTILPRSKI 157
 Db 178 SCVGIWILATVLSPELISD-LQSSSEQ-AMRSLTEHVE-AFITQVQMWIGFLV 234
 QY 158 ICLVWGLSVLIISSTFVNQKINTQGSVDCEPKYQTVTSEPRLWLLMLGL-ELLFGFFI 216
 Db 235 PLUMSCYLVIIITLLQARFNERNKAIVIAVVFVIFVOLFVGVVLAQTVANFNIT 294
 QY 217 PLMEIFCYFIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNNVLLV-TAANLGM 275
 Db 295 SSTCELSQNLNAYDVTVSLACVRCVNPFLVAFYIGVFRNDLFLKFDLGL 347
 QY 276 NRSQSEKLGITKTVTEVLAFLHCLNPLVYAFYIGQFRNRYFLKLDLWCV 328

RESULT 10
 ID IL8B_MACMU STANDARD; PRT; 353 AA.
 AC Q28519;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
 GN IL8RB OR CXCR2.
 OS MACACA MULATTA (RHESUS MACAQUE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96175151.
 RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
 RA LOPEZ-LARREA C.;
 RL IMMUNOGENETICS 43:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; X91116; E198176; -
 DR PROSITE; PS00237; G-PROTEIN_RECEPTOR; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT NON_TER 1 1
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 72 1 (POTENTIAL).
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 82 102 2 (POTENTIAL).
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 118 139 3 (POTENTIAL).
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 161 180 4 (POTENTIAL).
 FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 206 228 5 (POTENTIAL).
 FT DOMAIN 229 248 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 249 270 6 (POTENTIAL).
 FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 292 312 7 (POTENTIAL).
 FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 116 193 BY SIMILARITY.
 FT CARBOHYD 19 19
 FT NON_TER 353 353
 SQ SEQUENCE 353 AA; B592FE64 CRC32;

Query Match 34.5%; Score 959; DB 1; Length 353;
 Best Local Similarity 43.0%; Pred. No. 1.64e-163;
 Matches 129; Conservative 82; Mismatches 78; Indels 11; Gaps 8;

Db 34 APCRPESLEINKFYVYIYALVFLSLGNSLWMLVILSVGRSVTDVYLLNALADLL 93
 QY 27 APCRRSSSPG-YLYRIAYSLLICVLGLGNILVITFAFYKKARSMTDVLNMAIADIL 85

Db 94 FALTPLPIAAKNGV-WFGTEFLCKVYVLLKEVNFYSGILLACISVDYRLAIIVHATRTL 152
 QY 86 FVTLTFLPFAVSHATGAWVFSNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSF 145
 Db 153 TOK-RYL--VKFICLSIWGLSLLLALPVLFRRTVYSSNVPAC---YEDMGNNNTANRM 206
 QY 146 RLRSRTLPRSKIIICLVVWGLSVIISSTFVFNOK-YNTQGSVDCEPKYQTVSE-PIRWKL 203
 Db 207 LLRLIPQSFQFVPLLMFCYGFYTLTLFKAHMQKQHRAMRVIFAVVLPFLCWLDPYSL 266
 QY 204 LMLGLELLFGFFIPLMEIFCYFIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNM 263
 Db 267 VLLADTLMTQVIOETCERNHIDRALDATEILGILHCLNPLIYAFYIGQFRHGLLKIL 326
 QY 264 VLLVTAANLGMNR-SCQSEKLGITKTVTEVLAFLHCLNPLVYAFYIGQFRNRYFLKIL 322

RESULT 11
 ID CKR7_MOUSE STANDARD; PRT; 378 AA.
 AC P47774;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE C-C CHEMOKINE RECEPTOR TYPE 7 PRECURSOR (C-C CKR-7) (CCR7) (EBV-
 DE INDUCED G PROTEIN-COUPLED RECEPTOR 1) (EB11).
 GN CKBR7 OR EB11 OR EB1H.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6/CBA; TISSUE=THYMUS;
 RX MEDLINE; 95154835.
 RA SCHWEICKART V.L., RAPPORT C.J., GODISKA R., BYERS M.G., EDDY R.L. JR.,
 RA SHOWS T.B., GRAY P.W.;
 RL GENOMICS 23:643-650(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE ELC/MIP3B CHEMOKINE. PROBABLE MEDIATOR
 CC OF EBV EFFECTS ON B LYMPHOCYTES OR OF NORMAL LYMPHOCYTE FUNCTIONS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; L31580; G468341; -
 DR MGD; MGI:103011; CKBR7.
 DR PROSITE; PS00237; G-PROTEIN_RECEPTOR; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 24
 FT CHAIN 25 378 C-C CHEMOKINE RECEPTOR TYPE 7.
 FT DOMAIN 25 59 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 60 86 1 (POTENTIAL).
 FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 96 116 2 (POTENTIAL).
 FT DOMAIN 117 130 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 131 152 3 (POTENTIAL).
 FT DOMAIN 153 170 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 171 191 4 (POTENTIAL).
 FT DOMAIN 192 219 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 220 247 5 (POTENTIAL).
 FT DOMAIN 248 263 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 264 289 6 (POTENTIAL).
 FT DOMAIN 290 313 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 314 331 7 (POTENTIAL).
 FT DOMAIN 332 378 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 36 36
 FT DISULFID 129 210 BY SIMILARITY.
 SQ SEQUENCE 378 AA; 42941 MW; 9602A43B CRC32;

Query Match 34.5%; Score 959; DB 1; Length 378;
 Best Local Similarity 45.7%; Pred. No. 1.64e-163;
 Matches 134; Conservative 74; Mismatches 78; Indels 7; Gaps 6;

Db 60 FL-PLMYSVICFVGLGNGLVILTYIFKRLKMTDYLNLAVADILFLILPFWAYSE 118
 QY 38 YLYRIAYSLLICVLGLGNILVITFAFYKKARSMTDVLNMAIADILFLVLPFWAVSH 97

Db 119 AK-SWIFGVYCKIGFIYKLSFFSGMLLLCISIDRYAIVQAVSHRRHARVLLISKL 17
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Qy 98 ATGAWFVSNAQTCKLKGIYAINFENCGMLLTCTISMDRYIAIVQATKSRLRSNTLPRSKI 157
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 178 SCVGIMWLAFSLSPILLYSCLQNSGEDT--LRCSLVSAQVE-ALITIQVAQMVFGLV 234
I : . : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Qy 158 ICLVWGVLVSIISSSTFVNOKYNTQGSVDCEPKYQVPSEPIRWKLMLGL-ELLFGFFI 216
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 235 PMLAMSCFYLIIRTLQAQRNFERNKAIKVIIAAVVVFVFLPQNGVLAQVANFNIT 294
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Qy 217 PLMFMEICFYFIVKTLYQAQNSKEHKAIRVIAVVLFLACQIPHNNVLV-TAANLGKM 275
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 295 NSSCETSKQLNIADVITYSLASVRCCNPFLYAFIGVKFRSDLFKFKDGLCL 347
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Qy 276 NRSQSEKLGTYTKTVTEVLAFLHCCLNPLVYAFIGQKFRNYFLKILKDLCWCV 328
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 12
ID IL8B_PANTR STANDARD; PRS; 353 AA.
AC Q28807;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
GN IL8RB OR CXCR2.
OS PAN TROGLODYTES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN EUTHERIA; PRIMATES.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96175151.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROGES S.,
LOPEZ-LARREA C.;
RL IMMUNOGENETICS 43:261-267(1996).
CC - FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY.
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
CC SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; X91113; E198174; -
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT NON_TER 1 1
FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 72 1 (POTENTIAL).
FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 82 102 2 (POTENTIAL).
FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 180 4 (POTENTIAL).
FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 228 5 (POTENTIAL).
FT DOMAIN 229 248 6 (POTENTIAL).
FT TRANSMEM 249 270 7 (POTENTIAL).
FT DOMAIN 271 291 8 (POTENTIAL).
FT TRANSMEM 292 312 9 (POTENTIAL).
FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 19 19 POTENTIAL.
FT NON_TER 353 353
SQ SEQUENCE 353 AA; 39998 MW; F4564B58 CRC32;

Query Match 34.2%; Score 951; DB 1; Length 353;
Best Local Similarity 43.8%; Pred. No. 7.25e-162;
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps

Db 46 YFVVIYALVFLSLGNSLMVLNVSRGRSVTDVYLLNALADLLFALTPTWAASK 105
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

DR GCRDB: GCR_0077: -
 DR GCRDB: GCR_0610: -
 DR GCRDB: GCR_1001: -
 DR MW: 146928: -
 DR PROSITE: PS00237: G-PROTEIN_RECEPTOR; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 48 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 49 75 1 (POTENTIAL).
 FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 85 105 2 (POTENTIAL).
 FT DOMAIN 106 120 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 121 142 3 (POTENTIAL).
 FT DOMAIN 143 163 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 164 183 4 (POTENTIAL).
 FT DOMAIN 184 208 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 209 231 5 (POTENTIAL).
 FT DOMAIN 232 251 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 252 273 6 (POTENTIAL).
 FT DOMAIN 274 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 295 315 7 (POTENTIAL).
 FT DOMAIN 316 360 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 22 22 POTENTIAL.
 FT DISULFID 119 196 BY SIMILARITY.
 SQ SEQUENCE 360 AA; 40759 MW; 135CEAE CRC32;

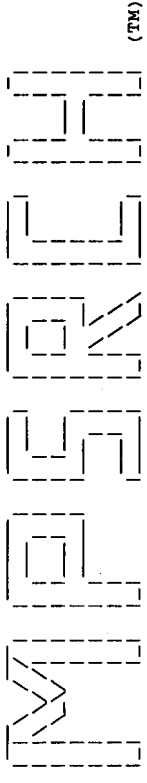
Query Match 34.2%; Score 951; DB 1; Length 360;
 Best Local Similarity 43.8%; Pred. No. 7.25e-162;
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;
 Db 49 YFVVIIVALLVLLSLGSLVMLVLYSRVGRSDVYLLNLALADLLFALTLPWAASK 108
 QY 38 YLYRIAYSLICVLGLGNLVITFAFKKARSDVYLLNLALADLLFALTLPWAASH 97
 Db 109 VNG-WIEGTFCKVSVLLKEVNFYSGILLACISVDRIYLAIVHATRLAQK-RYL--VRF 164
 QY 98 ATGAWFVSNAATCKLKGIVAINFNGMLLLTICISMDRYIAIVQATKSPRLSRTPRSKI 157
 Db 165 ICLSWGLSLLALPVLLFRRTVYSSNVSPAC---YEDMGNTANWRMLRLPQSGFTI 221
 QY 158 ICLVWGLSVIISSTFFVFNOK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215
 Db 222 VPLIMLCFCYFTLRTLFKAHMGOKHRAMRVIFAVLFLCWLPNVLLADTLMTQV 281
 QY 216 IPLMFIFCYFTFVKTQVQNSKRHRAIRVIAVVLVFLACQIPHNVLVTAANLGM 275
 Db 282 IQETCERNHIDRALDATEILGILHSCNLNPLIYAFIQGKFRGLLKIL 329
 QY 276 NR-SCQSEKLIGYKTVTEVLAFLHCLNPLVLYAFIQGKFRNYFLKIL 322

RESULT 14
 ID IL8B GORGO STANDARD; PRT; 353 AA.
 AC Q28422;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
 GN IL8RB OR CXCR2.
 OS GORILLA GORILLA GORILLA (LOWLAND GORILLA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96175151.
 RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
 RA LOPEZ-LARREA C.;
 RL IMMUNOGENETICS 43:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY.

CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: X91114; E198175; -
 KW PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT NON_TER 1 1
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 46 72 1 (POTENTIAL).
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 82 102 2 (POTENTIAL).
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 118 139 3 (POTENTIAL).
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 161 180 4 (POTENTIAL).
 FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 206 228 5 (POTENTIAL).
 FT DOMAIN 229 248 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 249 270 6 (POTENTIAL).
 FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 292 312 7 (POTENTIAL).
 FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 116 193 BY SIMILARITY.
 FT CARBOHYD 19 19 POTENTIAL.
 FT NON_TER 353 353
 SQ SEQUENCE 353 AA; 39919 MW; 4AF43313 CRC32;
 Query Match 34.0%; Score 945; DB 1; Length 353;
 Best Local Similarity 43.8%; Pred. No. 1.24e-160;
 Matches 126; Conservative 79; Mismatches 73; Indels 10; Gaps 7;
 Db 46 YFVVIIVALLVLLSLGSLVLYSRVGRSDVYLLNLALADLLFALTLPWAASK 105
 QY 38 YLYRIAYSLICVLGLGNLVITFAFKKARSDVYLLNLALADLLFALTLPWAASH 97
 Db 106 VNG-WIEGTFCKVSVLLKEVNFYSGILLACISVDRIYLAIVHATRLAQK-RYL--VRF 161
 QY 98 ATGAWFVSNAATCKLKGIVAINFNGMLLLTICISMDRYIAIVQATKSPRLSRTPRSKI 157
 Db 162 ICLSWGLSLLALPVLLFRRTVYSSNVSPVC---YEDMGNTANWRMLRLPQSGFTI 218
 QY 158 ICLVWGLSVIISSTFFVFNOK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215
 Db 219 VPLIMLCFCYFTLRTLFKAHMGOKHRAMRVIFAVLFLCWLPNVLLADTLMTQV 278
 QY 216 IPLMFIFCYFTFVKTQVQNSKRHRAIRVIAVVLVFLACQIPHNVLVTAANLGM 275
 Db 279 IQETCERNHINQALDATEILGILHSCNLNPLIYAFIQGKFRGLLKIL 326
 QY 276 NR-SCQSEKLIGYKTVTEVLAFLHCLNPLVLYAFIQGKFRNYFLKIL 322

RESULT 15
 ID CKR4 HUMAN STANDARD; PRT; 360 AA.
 AC P51679;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE C-C CHEMOKINE RECEPTOR TYPE 4 (C-C CKR-4) (CXCR-4) (K5-5).
 GN CKBR4.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 95370289.
 RA POWER C.A., MEYER A., NEMETH K., BACON K.B., HOOGWERF A.J.,
 RA PROUDFOOT A.E.I., WELLS T.N.C.;
 RL J. BIOL. CHEM. 270:19495-19500(1995).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA
 CC AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY INCREASING THE
 CC INTRACELLULAR CALCIUM IONS LEVEL.

Search completed: Fri Nov 13 12:05:50 1998
Job time : 40 secs.



(TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 12:06:08 1998; MasPar time 21.91 Seconds
Tabular output not generated. 829.627 Million cell updates/sec

Title: >US-08-887-977-10
Description: (1-365) from US08887977.pep
Perfect Score: 2779
Sequence: 1 MFSTPVKILLCQSILHITQL.....NISQTSATDNDNASSFTM 365

Scoring table:
PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb16
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 47.707; Variance 110.430; scale 0.432

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	2051	73.8	367	11	054689 G PROTEIN-COUPLED RECE	0.00e+00
2	833	30.0	368	13	042444 INTERLEUKIN-8-LIKE REC	1.60e-119
3	824	29.7	359	11	055169 RECEPTOR PROTEIN CCR3	6.32e-118
4	820	29.5	415	4	015185 G PROTEIN-COUPLED RECE	3.24e-117
5	817	29.4	359	11	054814 CHEMOKINE RECEPTOR CCR	1.10e-116
6	799	28.8	352	6	062743 CHEMOKINE RECEPTOR CCR	1.70e-113
7	799	28.8	352	6	062746 CHEMOKINE RECEPTOR CCR	1.70e-113
8	798	28.7	352	6	062744 CHEMOKINE RECEPTOR CCR	2.55e-113
9	796	28.6	352	6	062745 CHEMOKINE RECEPTOR CCR	5.76e-113
10	786	28.3	352	6	018770 CCR5 RECEPTOR (FRAGMEN	3.39e-111
11	776	27.9	352	4	015538 CCR5 RECEPTOR (FRAGMEN	1.98e-109
12	772	27.8	373	11	055193 CHEMOKINE RECEPTOR CCR	1.01e-108
13	769	27.7	352	6	018771 CCR5 RECEPTOR (FRAGMEN	3.42e-108
14	769	27.7	352	6	018772 CCR5 RECEPTOR (FRAGMEN	3.42e-108
15	766	27.6	384	4	000537 CHEMOKINE RECEPTOR CCR	1.16e-107
16	763	27.5	384	4	000590 CG-CHEMOKINE RECEPTOR	3.92e-107
17	756	27.2	360	6	018793 CHEMOKINE RECEPTOR	6.73e-106
18	739	26.6	333	4	014694 CCR5 RECEPTOR (FRAGMEN	6.67e-103
19	734	26.4	378	11	008707 CHEMOKINE (C-C) RECEPT	5.07e-102
20	722	26.0	382	11	009027 CCR10-RELATED RECEPTOR	6.55e-100

21	717	25.8	360	4	060835 CXCR4 GENE ENCODING RE	4.36e-99
22	710	25.5	352	6	046428 ALPHA-CHEMOKINE RECEPT	8.43e-98
23	708	25.5	352	6	062747 CHEMOKINE RECEPTOR CX	1.89e-97
24	683	24.6	383	14	Q89609 G PROTEIN-COUPLED RECE	4.50e-93
25	513	18.5	353	13	P79960 MESENCHYME-ASSOCIATED	1.35e-63
26	506	18.2	372	11	O70526 BRADYKININ B2 RECEPTOR	2.13e-62
27	482	17.3	332	11	O62973 CHEMOKINE RECEPTOR LCR	2.62e-58
28	482	17.3	330	4	O15132 P2Y5-LIKE RECEPTOR	2.62e-58
29	460	16.6	383	13	O42324 MU-OPIOID RECEPTOR	1.41e-54
30	457	16.4	238	13	Q92158 ANGIOTENSIN II RECEPTO	4.55e-54
31	449	16.2	185	13	O42445 CX3 CHEMOKINE RECEPTOR	1.02e-52
32	441	15.9	381	13	O42402 ORNITHOKININ RECEPTOR	2.28e-51
33	438	15.8	361	11	O35811 G-PROTEIN COUPLED RECE	7.30e-51
34	437	15.7	373	13	O57585 OPIOID RECEPTOR HOMOLO	1.08e-50
35	436	15.7	404	4	O35218 G-PROTEIN COUPLED RECE	1.58e-50
36	421	15.1	395	11	Q84166 ADRENOMEDULLIN RECEPTO	5.22e-48
37	415	14.9	303	13	P70058 ANGIOTENSIN RECEPTOR R	5.27e-47
38	414	14.9	374	13	O57466 G PROTEIN COUPLED P2Y	7.75e-47
39	394	14.2	344	4	O00421 CCR6	1.67e-43
40	390	14.0	345	11	O70129 ANAPHYLATOXIN C5A RECE	7.70e-43
41	375	13.5	375	4	O43494 G PROTEIN-COUPLED RECE	2.34e-40
42	359	12.9	334	11	Q61125 BRADYKININ RECEPTOR, B	1.01e-37
43	351	12.6	168	11	O35715 SOMATOSTATIN RECEPTOR	2.06e-36
44	336	12.1	361	6	O46685 ORPHAN G PROTEIN-COUPLE	5.73e-34
45	329	11.8	360	11	O70171 L-CCR	7.83e-33

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	367 AA.
ID	054689			
AC	054689			
DT	01-JUN-1998 (TREMBLREL. 06, CREATED)			
DT	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)			
DT	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)			
DE	G PROTEIN-COUPLED RECEPTOR KV411.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
NC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	YANAGIHARA S., KOMURA E., YAMAGUCHI Y.,			
RL	SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; AB009369; D1024681; -			
SQ	SEQUENCE 367 AA; 42102 MW; 64D91F19 CRC32;			
Query Match	73.8%;	Score 2051;	DB 11;	Length 367;
Best Local Similarity	78.8%;	Pred. No. 0.00e+00;		
Matches	256;	Conservative 43;	Mismatches 25;	Indels 1; Gaps 1;

Db	43	IAYSLLICVGLGIMVMTFAFYKARSMTDVLNNMAITDILFVLTLPFWAVTHATNT 102
QY	42	IAYSLLICVGLGILVVITFAFYKARSMTDVLNNMAIADILFVLTLPFWAVSHATGA 101
Db	103	WVFSALCKLMKGYANFNCGMLLACISMDRYIAIQAATKSFVRSTRTTHSKVICVA 162
QY	102	WVFSNATCKLKGIAINFNCGMLLTATCISMDRYIAIQAATKSFRLSRITLPRSKII 161
Db	163	VWFSIIISSTPTFNKKYELQDRDCEPRYSRVSSEPTWKLGLGELFGFTPLIFM 222
QY	162	VWGLSVIIISSTPTFNKKYELQDRDCEPRYSRVSSEPTWKLGLGELFGFTPLIFM 221
Db	223	VFCVLIKTIVQAQNSKRRAIRVIAVLVAVLVLACQIPHNVLVLTAVNTGKVRSCST 282
QY	222	IFCTVFIKTLVQAQNSKRRAIRVIAVLVAVLVLACQIPHNVLVLTAVNTGKVRSCQ 281
Db	283	EKVLATRNVAEVLAFHCCCLNPLYAFIGOKFRNYFNKIMKDVCMRRKMKMFGFLCAR 342
QY	282	EKLGYKTKTEVLAFHCCCLNPLYAFIGOKFRNYFNKIMKDVCMRRKMKMFGFLCAR 341
Db	343	YISESYISROTSETVENDNASSFTM 367
QY	342	RYSEN-ISROTSETVENDNASSFTM 365

Qy	103	VFSNATCKLLGVIYAINFNGMILLTCTISMDRYIAIVQATKSFRLSRTRLPSRKIIICLV	162
Db	162	WGLAVLAALPEFIFHESQDNFGDLSCSPRYPE-GEEDSKWRKSHALRMNT-FGLALPLLIM	219
Qy	163	WGLSVIISSTTFVFNQYNTQSGDCEPKQTVSEPIRNK-LMLGLELLFGFFPLPMEM	221
Db	220	VICYSGIIKTLRLCPNKKKHKAQLIFVMMVEFFITWTPYLNLLLSAFHSFTLETSCQ	279
Qy	222	IFCYTFIVKTLVQAQNSKRHKAIRVIAVYVLACQIPHNNVLLVTAANLGMNRSCQS	281
Db	280	SIHDLAMQVTEVITHTHCCINPIIYAFVGERFRKH-LRLF	319
Qy	282	EKLGYKTYTTEVLAFHLLHCLNPLVYAFIQGKERNFYFLKIL	322

```

RESULT      4
ID          015185
AC          PRELIMINARY;      PRT;      415 AA.
AD          015185;
DT          01-JAN-1998 (TREMBLREL. 05, CREATED)
DT          01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT          01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE          G PROTEIN-COUPLED RECEPTOR CKR-L2.
DE          OS HOMO SAPIENS (HUMAN).
OS          EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC          EUTHERIA; PRIMATES.
RN          [1]
RP          SEQUENCE FROM N.A.
RA          GUTIERREZ J.; VARONA R.; ZABALLOS A.; LIND P.; MARQUEZ G.;
RL          SUBMITTED (SEP-1996) TO EMBL/GENBANK/DDSI DATA BANKS.
DR          EMBL; Z79783; E264773; -.
DR          PFAM; PF00001; 7tm.1
SQ          SEQUENCE 415 AA; 45608 MW; EB0DF045 CRC32;

```

RESULT	5	
ID	054814	PRELIMINARY; PRT; 359 AA.
AC	054814;	
DT	01-JUN-1998	(TREMBLREL. 06, CREATED)
DT	01-JUN-1998	(TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT	01-JUN-1998	(TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE	DE	CHEMOKINE RECEPTOR CCR3.
OS	OS	RATTUS NORVEGICUS (RAT).
OC	OC	EUKARYOTA; METAQOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
EN	EN	EUTHERIA; RODENTIA.
		[1]
RP	RP	SEQUENCE FROM N.A.
RC	RC	STRAIN-WISTAR; TISSUE-SPLEEN;
RA	RA	JTANG Y., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L., SONNT
RA	RA	DEFIEBRE C.M., PENNELL N.A., STREIT W.J., HARRISON J.K.;

RESULT	2	
ID	042444	PRELIMINARY; PRT; 368 AA.
AC	042444;	
DT	01-JAN-1998 (TREMREL. 05, CREATED)	
DT	01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)	
DT	01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)	
DE	INTERLEUKIN-8-LIKE RECEPTOR.	
OS	ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;	
OC	EUSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.	
RN	[1]	

```

RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR EMBL; AU000159; E1193478; -.
PFAM; PF00001; 7tm_1.
SQ SEQUENCE 368 AA; 41523 MW; 56D3903B CRC32;

Query Match          30.0%; Score 833; DB 13; Length 368;
Best Local Similarity 41.1%; Pred. No. 1.60e-119;
Matches 116; Conservative 76; Mismatches 86; Indels 4; Gaps 4;

Db      54 YNSIVILGGIGNTVYVIYLHFRQRLLKMTDTIYLLNLAVADLFGLTLEPLWAF-EANQGW 112
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      44 YSLICVLGILLNLIYF-ITFAFYKKARSMTDVYLLNNAIDILVTLFPFWAVSHATGAW 102
        | : : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      113 SFLGLCKVTSAFYKINFFSSMLLTCTISVDRVVVIVQTMAQNSKRQLSCSKFCVCAY 172
        | : : | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      103 VFSNATCKLLKGIAINFNCGMLLLLTCISMDDRYIAIVQAITSFRSRRTLPKRSKIICLV 162

```

[illegible]

	Query Match	29.7%	Score 824;	DB 11;	Length 359;
	Best Local Similarity	37.7%;	Pred. NO. 6.32e-113;		
	Matches 106;	Conservative 91;	Mismatches 76;	Indels 8;	Gaps 7;
Db	45 YSLVFTVGLLGNMVMYLIILKYRKQIMTYILLNLNLAISDLFLFETPFW--IHYVLWNEW 103 : : : :: : : : : : : : : : : : : : : : : :				
QY	44 YSLICVGLLGNLYIVITFAFKKARSMTDVYLNNMAIDILVFUTLPFWVSHAT--GAW 102 : : : : :				
Db	104 GFCHCMCKMLSGLYIALYSEIFFILLTRIDRYLATVHAV--LALRARTVTPTATITSIT 161 : :				

RL J. NEUROIMMUNOL. 0:0-0(1998).
 DR EMBL; AF003954; G2897073; -.
 SQ SEQUENCE 359 AA; 41643 MW; CLFC70CA CRC32;

Query Match 29.4%; Score 817; DB 11; Length 359;
 Best Local Similarity 37.4%; Pred. No. 1.10e-116;
 Matches 105; Conservative 92; Mismatches 76; Indels 8; Gaps 7;

Db	45	YSLVFIYGLLGNNMMVLILIKRYKQLQIMTNIYLNLNAISDLLEFLFVPFW--IHYVLWNEW	103
Qy	44	YSLICVLGLGNILVITFAFYKKARSMTDYYLLNMAIDILFVLTPFWAVSHAT-GAW	102
Db	104	GFGHCCKMLSGLYLYALYSEIFFILLTTDRILAIVHAV--IALRARTVTFTTSITT	161
Qy	103	VFSNATCKLKGIYAINFNGCMLLTCTISMDRYIAIQVATKSFLRSRTLPKSKIIICLVV	162
Db	162	WGFAVLAALPEEFIFHESODNFGLSCSPRYPE-GEEDSMKRPHALARINI-FGLALPLLIM	219
Qy	163	WGLSVIISSTTFVNQKYNTQGSDVCBPKQTVPSEPIRWK-LMLGLELFFGFFIPLMFM	221
Db	220	VICYSGIITKLLRCNPKNKHKAIOLIPIVMIVFFIPTYNLVLLLSAFESTPLETSCQQ	279
Qy	222	IFCYTFIVKTIVQAQNSKRHKARVIIVAVLVFLACQIPHNVLLTYAANLGRKNRSCQS	281
Db	280	SIHLDLAMQVTEVITHPHCCINPIIYAFVGPERFKH-LRLF	319
Qy	282	EKLIGYTKTVTEVLAFLHCLLNPLVYAFIGOKFRNYFLKIL	322

RESULT 6 PRELIMINARY; PRT; 352 AA.

ID	O62743	AC
DT	01-AUG-1998	(TREMBLREL. 07, CREATED)
DT	01-AUG-1998	(TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT	01-AUG-1998	(TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE	CHEMOKINE RECEPTOR CTYS.	
OS	CERCOCABUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
RN	EUTHERIA; PRIMATES.	[1]
RP	SEQUENCE FROM N.A.	
RA	CHEN Z., GETTIE A., HO D.D., MARK P.A.;	
RL	VIROLOGY 0:0-0(1998).	
DR	EMBL; AF051902; G3135296; -.	
SQ	SEQUENCE 352 AA; 40407 MW; 1716CC5C CRC32;	

Query Match 28.8%; Score 799; DB 6; Length 352;
 Best Local Similarity 38.9%; Pred. No. 1.70e-113;
 Matches 109; Conservative 88; Mismatches 72; Indels 11; Gaps 10;

Db	37	YSLVFIYFGVGNILVILINKCRKLSMTDIYLLNLAISDLLEFLFVPFWA-HYAAQWD	95
Qy	44	YSLICVLGLGNILVITFAFYKKARSMTDYYLLNMAIDILFVLTPFWAVSHATGAW	103
Db	96	FGNTKCOLLTGLYFIGFGSIFITLLTDRIYLAIVHA--GFALKARTVTFGVVTSVITW	153
Qy	104	FSNATCKLLKGIAINENCMLLTCTISMDRYIAIQVATKSFLRSRTLPKSKIIICLVV	163
Db	154	VWVFASLPGIIFTRS-QRGELHYTCSPHPYEQYF-WKNFOTLKIVL-GLVLPILVYM	210
Qy	164	GLSVIISSTTFVNQKYNTQGD-VCEPKYTQVSEPIRWK-LMLGLELFFGFFIPLMFM	221
Db	211	VICYSGIKTLKLRCKNEKKHRAVRLLFTIMIVYELFWAPYNTVLLANTFOEPFGLN-NC	269
Qy	222	IFCYTFIVKTIVQAQNSK-RHKAIRVIIVAVLVFLACQIPHNVLLTYA-ANLGRKNRSC	279
Db	270	SSNRLDQAMQVTEITGLTMHCCINPIIYAFVGKFRNYLL	309
Qy	280	QSEKLGITYKTVTEVLAFLHCLLNPLVYAFIGOKFRNYFL	319

RESULT 7 PRELIMINARY; PRT; 352 AA.

AC	O62745;
DT	01-AUG-1998 (TREMBLREL. 07, CREATED)
DT	01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT	01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE	CHEMOKINE RECEPTOR CCR5.
OS	CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; PRIMATES.
[1]	
RN	SEQUENCE FROM N.A.
RA	CHEN Z., GETTIE A., HO D.D., MARX P.A.;
RL	VIROLOGY 0:0-0(1998).
DR	EMBL; AF051905; G3135302; -.
SQ	SEQUENCE 352 AA; 40489 MW; FB9CE731 CRC32;
 Query Match 28.8%; Score 799; DB 6; Length 352; Best Local Similarity 38.9%; Pred. No. 1.70e-113; Matches 109; Conservative 87; Mismatches 73; Indels 11; Gaps	
Db	37 YSLVFIQFVGNIILVLLINCKRLKSMTDIYLNLAIISDLLELTLPFWA-HYAAQWD 95 : : : : : : : : : : : : : : :
Qy	44 YSLCVGLGNLVITVFAYKKARSMTDVYLNMAADILVLTLFPWAVSHATGAW 103 : : : : : : : : : : : : : : :
Db	96 FGMTMCOLLGWTFIFGFSGIFFILLIDRYLAIVHAV--FALKARTVTFLGVTSVITW 153 : : : : : : : : : : : : : : :
Qy	104 FSNATCKLLKGIVAINFCNGMLLTCSIMSDRYIAIQVATKSFRLSPRSKIICLVW 163 : : : : : : : : : : : : : : :
Db	154 VVAVFASLPFGIIFTRS-QREGLHYTCSPHPYSQYQF-WKNFOTIKIVIL-GLVPLLYM 210 : : : : : : : : : : : : : : :
Qy	164 GLSVIISSTSEFNQYNQTGSD-VCEPKYQTVSEPIRWK-LLMGLELLFGFFIPLEM 279 : : : : : : : : : : : : : : :
Db	211 VICVSGILTKLRGNEKKRHRVRLPFTIMIVLFWPAPNVIVLLNTFOEFGLN-NC 269 : : : : : : : : : : : : : : :
Qy	222 IFCTYFIVKTLLVQAQNSK-RHKAIRVIITAVVLFLACQIPHNMVLLVTA-ANLGKMRSC 279 : : : : : : : : : : : : : : :
Db	270 SSSNRLDQAMOVTTGLMTHCCINPLIYAFVGEKFRNYLL 309 : : : : : : : : : : : : : : :
Qy	280 QSEKLIGTKIVTEVLAFRLCCLNPVLYAFIQGFPRNFYL 319 : : : : : : : : : : : : : : :
RESULT	8 PRELIMINARY; PRT; 352 AA.
ID	O62745
AC	O62745; PRELIMINARY; PRT; 352 AA.
DT	01-AUG-1998 (TREMBLREL. 07, CREATED)
DT	01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE	CHEMOKINE RECEPTOR CCR5.
OS	CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; PRIMATES.
[1]	
RN	SEQUENCE FROM N.A.
RA	CHEN Z., GETTIE A., HO D.D., MARX P.A.;
RL	VIROLOGY 0:0-0(1998).
DR	EMBL; AF051904; G3135300; -.
SQ	SEQUENCE 352 AA; 40460 MW; E6A5AA07 CRC32;
 Query Match 28.7%; Score 798; DB 6; Length 352; Best Local Similarity 37.9%; Pred. No. 2.55e-113; Matches 113; Conservative 92; Mismatches 80; Indels 13; Gaps	
Db	19 PCQRINGQIAARLLPPYLSLVFIFFGVNLIWLINCKRLKSMTDIYLNLAIISDLL 78 : : : : : : : : : : : : : : :
Qy	28 PCRR-SGSS-PGYLYRTAYSILCVLLGNLVITVFAYFKARSMTDVLNLNAADIL 85 : : : : : : : : : : : : : : :
Db	79 FLTLVPFWA-HYAAAQWDFGNTMCOLLTGLYFIFGFSGIFFILLIDRYLAIVHAV--F 135 : : : : : : : : : : : : : : :
Qy	86 FVLTLFPWAVSHATGAWWESNATCKLLKGIVAINFCNGMLLTCSIMSDRYIAIQVATKSF 145 : : : : : : : : : : : : : : :
Db	136 ALKARTVFGVTVSITWVAVFASLPGIIFTRS-QREGLHYTCSPHPYSQYQF-WKNE 193 : : : : : : : : : : : : : : :
Qy	146 RLRSRTLPRSKIICLVVWGSLVSISTSEFNQYNQTGSD-VCEPKYQTVSEPIRWK-L 203 : : : : : : : : : : : : : : :

```

Db 194 QTLKIVIL-GUVLPPLVWVICYSGILKTLRCRNEKRHRAVLFTIMTVYFLWAPYN 252
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 204 LMLGLELLFGFFIPLMFIMFYCTEIVTKLVAQNSK-RKHAIRVIAVLVLACQIPHN 262
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 253 IVLLNLTFFQFFGLN-NCSSNRLDQAMQVTEILGHTWTHCCINFIITAFVCEKERNVLL 309
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 263 MVLVTA-ANLGRKNRSCQSEKLGYTKTVEVLAFLHCLNPLVYAFIGQKFRNYFL 319

RESULT 9
ID 062744 PRELIMINARY; PRT; 352 AA.
AC 062744;
DT 01-AUG-1998 (TREMBUREL 07, CREATED)
DT 01-AUG-1998 (TREMBUREL 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBUREL 07, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR5.
OS CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RP CHEN Z, GETTIE A, HO D.D., MARX P.A.;
RL VIROLOGY 0:0-(1996).
DR EMBL; AF051903; G3135298; -.
SQ SEQUENCE 352 AA; 440503 MW; C922372D CRC32;

```

Query Match	28.6%	Score 796;	DB 6;	Length 352;
Best Local Similarity	38.9%;	pred. No. 5.76e-113;		
Matches	109;	Conservative	87;	Mismatches 73; Indels 11; Gaps 10;
<hr/>				
Db	37	YSLVFTEFGVGNILVLVLLINCRLKLSMTDIYLNLNAISDLLFLLVPPFWA-HYAAQWD	95	
		: : : : : : :		
QY	44	YSLCVLGLLGNIWLVTFAFYKKARSMTDVIYNMAIDILEVLTLPFWAVSHATGAW	103	
<hr/>				
Db	96	FGNTMCOQLLGLVFI GFESGIFFILLTIDRYLAIVHAV--FALKARTVTFGLVTSVIITW	153	
		: : : : : : : : : : : : : :		
QY	104	FSNATCKLLGGIYAINPNCGMLLTCTISMDRYTAIVQATSFRLRSRTLPKRKIICLVVM	163	
<hr/>				
Db	154	VVAVFASLPGIIITRS-QRSGLHWTCSPHPYPYSOYOF-WKNFTKLKIVIL-GVLVPLLVAM	210	
		: : : : : : : : : : : : : : : : : :		
QY	164	GLSVIISSTVFVNQKNTOGSD-VCEPKQTVPSEPIRWK-LMLGLELUGFFIFLPMFM	221	
<hr/>				
Db	211	VICYSGIKTLTLRCNEKKRHRAVRILFTIMIFYFWAPYNIVLLINTTFQEEFGLN-NC	269	
		: : : : : : : : : : : : : : : : : : : : :		
QY	222	IFCYTFVKTLVQAQNSK-RHKAIRVIIAVLVFLACQIPHNNVLLVTA-ANLGKNRSC	279	
<hr/>				
Db	270	SSNRLDQAMQVTTETLGMTCCCNIPPIIAFYVGKEFRNYLL	309	
		: : : : : : : : : : :		
QY	280	QSKELGYTTFVTVEVLAFLCCLNPVLYAFIGOKFRNYFL	319	

RESULT	10
ID	O18770 PRELIMINARY; PRT; 352 AA.
AC	O18770;
OC	O18770; 05, CREATED)
DT	01-JAN-1998 (TREMBREL. 05,
DT	01-JUN-1998 (TREMBREL. 06, LAST ANNOTATION UPDATE)
DE	CCRS RECEPTOR (FRAGMENT).
DN	CCRS.
OS	PAN TROGLODYTES (CHIMPANZEE).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; PRIMATES.
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=MACCRS-140A:
RC	ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG G., HAHN B., HO D.D.;
RA	HOMID.
RL	AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).
DR	EMBL; AF011538; G2305194; -.
DR	PFAM; PF00001; 7tm_1.
FT	NON_TER 352
SQ	SEQUENCE 352 AA; 40523 MW; FLC10E99 CRC32;

Query Match	28.3%;	Score 786;	DB 6;	Length 352;	
Best Local Similarity	38.2%;	Pred. No. 3.39e-111;			
Matches	107;	Conservative	87;	Mismatches 75;	Indels 11; Gaps 10

Db	37	YSLVFIFGFGVNILVWLINCKRLKSMTDIYLLNLAISDLFLUTVPFWA-HYAAAQWD	95
Qy	44	YSLCVLGLLNILWITFAFYKKARSMTDYLLNMAIADILFLVTLTFWAVSHATGAW	103
Db	96	FGNTMCCOLLGLYFIFGFGFIILLTIDRYLAIVHAV--FALKARTVTGCVTSVITW	153
Qy	104	FSNATCKLKGIAINFNCGMLLTCISMDRYAIVQTKFSRSLRTPSKKIICLVVW	163
Db	154	VVAYFASLPGIIFTRS-QREGLHYTCSSHFPYSQYF-WKFNQTLKMWIL-GLVLPPLVM	210
Qy	164	GLSVIISSTSEPNQYNTQSD-VCEPKYQTVSEPIRWK-LLMLGLELLFGFFIPLMF	221
Db	211	VICYSGLIKTLRCNEKKRRRAVRLFTIMIVELFWAPNIVLLNFTFOEFTGLNCS	270
Qy	222	IFCYTFIVKTLVQANSK-RHKAIRVIAVVLVFLACQIPHNMVLLVTA-ANLGKMNRS	279
Db	271	SCNRL-DQAMOVTELTGLNTHCCINPIIVAFYGEKFRNLL	309
Qy	280	QSEKLGYTKVTEVLAFHCLNPLVYAFIQGRNFYL	319

RESULT	11	
ID	O15538	PRELIMINARY; PRT; 352 AA.
AC	O15538;	
DT	01-JAN-1998	(TREMBREL. 05, CREATED)
DT	01-JAN-1998	(TREMBREL. 05, LAST SEQUENCE UPDATE)
DT	01-JUN-1998	(TREMBREL. 06, LAST ANNOTATION UPDATE)
DE	CCR5 RECEPTOR (FRAGMENT).	
GN	CCR5.	
OS	HOMO SAPIENS (HUMAN).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	EUTHERIA; PRIMATES.	
RC	[1]	
RP	SEQUENCE FROM N.A.	
RA	ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG G., HAHN B.,	
RA	HO D.D.;	
RL	AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).	
DR	EMBL; AF011536; G2305190; -.	
DR	EMBL; AF011516; G2305150; -.	
DR	EMBL; AF011534; G2305186; -.	
DR	PFAM; PF00001; 7tm.1.	
FT	NON_TER	352
FT	SEQUENCE	352 AA; 40552 MW; BLECA8D9 CRC32;

Query Match	27.9%;	Score 776;	DB 4;	Length 352;	
Best Local Similarity	37.1%;	Pred. No. 1.98e-109;			
Matches	104;	Conservative	92;	Mismatches 73;	Indels 11; Gaps 10

[illegible]

```

RESULT 12
ID O55193 PRELIMINARY; PRT; 373 AA.
AC O55193;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY;
RA JIANG Y., SALAFRANCA M.M., ADHIKARI S., XIA Y., FENG L., SONNTAG M.K.,
RA DEFEIERE C.M., PENNELL N.A., STREIT W.J., HARRISON J.K.;
RL J. NEUROIMMUNOL. 0:0-0(1998).
DR EMBL; U77349; G2896818; -.
SQ SEQUENCE 373 AA; 42763 MW; 14578A08 CRC32;

Query Match 27.8%; Score 772; DB 11; Length 373;
Best Local Similarity 37.7%; Pred. No. 1.0le-108;
Matches 113; Conservative 90; Mismatches 81; Indels 16; Gaps 14;

Db 44 PCHKTSVQIGAWLPPPLYSLVFIFGVGNMLVLIILISCKKLKMTDYLFLNLAISDLL 103
QY 28 PCRRSGSSP-G-YLRYIASLVGLGNILVITFAFYKKARSMTDYLNNATADIL 85

Db 104 FLLTLPFWA-HYAANEWFNGTCKLFTGLYHIGYFGGIFILLTIDRYLAIVHAV--F 160
QY 86 FVLTLFPWAVSHATGAWVFSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATKSF 145

Db 161 ALKARTVTFGVITVTVVWVAVFASLPGIIFTKSEDDQHTGCP-Y--FP-TI-WKNFQ 215
QY 146 RLRSRTLPKRSKILICLVWGLSVIISSTFVNQKYNQTSQDVCEPKYQTVSEPIRWKLLM 205

Db 216 TIMRNILSILPLVWVICYSGLHTLFCRNEKKRHRAVRLIFAIMIVYFLEWTPVNIY 275
QY 206 LGLELFGFFLPMFIMFYCTIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVW 264

Db 276 LFLATTFQFBLG-MS-NCVDMHLDQAMQVETLGMTHCCVNPYIYAFVGEKFRY-LSIF 332
QY 265 LLVTA-AN-LGKMNRSQCEKLGITKTVTEVLAFHLHCLNPLVYAFIGQKFRNYFLKIL 322

RESULT 13
ID O18772 PRELIMINARY; PRT; 352 AA.
AC O18772;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE CCR5 RECEPTOR (FRAGMENT).
GN CCR5.
OS PAN TROGLODYTES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHCR5-142A;
RA ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG G., HAHN B.,
RA HO D.D.;
RL AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).
DR EMBL; AF011541; G2305200; -.
DR PFAM; PF00001; 7tm1.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40598 MW; A9BF8EDF CRC32;

Query Match 27.7%; Score 769; DB 6; Length 352;
Best Local Similarity 36.8%; Pred. No. 3.42e-108;
Matches 103; Conservative 92; Mismatches 74; Indels 11; Gaps 10;

Db 37 YSLVIFGVGNMLVILILINCKRLKSMTDIYLLNLAISNLFLLITVPFWA-HYAAAQWD 95
QY 44 YSLICVLGLGNILVITFAFYKKARSMTDYLNNMAIADILFVLTPFWAVSHATGAW 103

Db 96 FGNTMCOLLGLYFIFGSGIFFILLTIDRYLAIVHAV--FALKARTVTFGVTSVITW 153
QY 104 FSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKRSKILCVW 163

Db 154 VVAVFASLPGIIFTRS-QKEGLHYTCSHFPYSQOF-WKNFQTLKIVIL-GLVPLLVW 211
QY 164 GLSVIISSTFVNQKYNQTSQDVCEPKYQTVSEPIRWK-LMLGLELFGFFLPMFIM 222

Db 212 ICYSGILKTLRCRNEKKRHRAVRLIFAIMIVYFLEWTPVNIYAFVGEKFRNYLL 309
QY 223 FCTYFIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVLLVTA-ANLGMNRSQ 280

Db 271 SSNRLDQAMQVETLGMTHCCINPIIYAFVGEKFRNYLL 309
QY 281 SEKLIGYTKTVTEVLAFHLHCLNPLVYAFIGQKFRNYFL 319

Query Match 27.7%; Score 769; DB 6; Length 352;
Best Local Similarity 36.8%; Pred. No. 3.42e-108;
Matches 103; Conservative 92; Mismatches 75; Indels 9; Gaps 8;

Db 37 YSLVIFGVGNMLVILILINCKRLKSMTDIYLLNLAISNLFLLITVPFWA-HYAAAQWD 95
QY 44 YSLICVLGLGNILVITFAFYKKARSMTDYLNNMAIADILFVLTPFWAVSHATGAW 103

Db 96 FGNTMCOLLGLYFIFGSGIFFILLTIDRYLAIVHAV--FALKARTVTFGVTSVITW 153
QY 104 FSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKRSKILCVW 163

Db 154 VVAVFASLPGIIFTRS-QKEGLHYTCSHFPYSQOF-WKNFQTLKIVIL-GLVPLLVW 211
QY 164 GLSVIISSTFVNQKYNQTSQDVCEPKYQTVSEPIRWK-LMLGLELFGFFLPMFIM 222

Db 212 ICYSGILKTLRCRNEKKRHRAVRLIFAIMIVYFLEWTPVNIYAFVGEKFRNYLL 309
QY 223 FCTYFIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVLLVTA-ANLGMNRSQ 280

Db 271 SSNRLDQAMQVETLGMTHCCINPIIYAFVGEKFRNYLL 309
QY 281 SEKLIGYTKTVTEVLAFHLHCLNPLVYAFIGQKFRNYFL 319

RESULT 15
ID O00537 PRELIMINARY; PRT; 384 AA.
AC O00537;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR-9.
OS HOMO SAPIENS (HUMAN).

```

CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/410.454A
 CC FILING DATE: 24-Mar-1995
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/234494
 CC FILING DATE: 28-APR-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/677211
 CC FILING DATE: 29-MAR-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Love, Richard B.
 CC REGISTRATION NUMBER: 34,659
 CC REFERENCE/DOCKET NUMBER: P0706C1D3
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415/225-5530
 CC TELEFAX: 415/952-9881
 CC TELEX: 910/371-7168
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 350 amino acids
 CC TYPE: Amino Acid
 CC TOPOLOGY: Linear
 CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 1; Length 350;
 Best Local Similarity 44.4%; Pred. No. 8.26e-79;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMELETLNKKVYIIAYALVLLSLGNSLVMLVILYRGVRSVTDVYLLNLALAD 85
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNILVITFAFYKARSDTDVYLLNMAID 83
 Db 86 LLFALTLPWAASKVNG-WIFGTFCKVSVLKEVNFYSGILLACISVDRIYLAIVHATR 144
 QY 84 ILFVLTLPFWAVSHATGAWFSNATCKLLGIYAINFCNCGMLLTCTISMDRYIAIVQATK 143
 Db 145 TLTKQ-RHL--VKFVLCGCGWLSMNLSPFLFRQAYHPNNSPVC---YEVIGNDTAKW 198
 QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
 Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTTLRTLFKAHMGQKHRAVRVFAVVLFLCWLPL 258
 QY 202 KLLMLGELLGFFPLMFIMFCYFIVKTLVQAQNSKRKRAIRVIAVVLFLACQIPH 261
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCLNPIIYAFIQGNFRHGFLK 318
 QY 262 NWLLVTAANLGMKMR-SCQSEKLGITKTVTEVLAFLHCLLPVLYAFIQGNFRNYFLK 320
 Db 319 IL 320
 QY 321 IL 322

RESULT 4
 ID US-08-076-093A-2 STANDARD; PRT; 350 AA.
 XX
 AC xxxxxx
 CC
 CC

Sequence 2, Application US/08076093A
 Sequence 2, Application US/08076093A
 Patent No. 5543503
 GENERAL INFORMATION:
 CC APPLICANT: Chuntharapai, Anan
 CC APPLICANT: Lee, James
 CC APPLICANT: Hebert, Caroline
 CC APPLICANT: Jin Kim, K.
 CC TITLE OF INVENTION: Antibodies to Human PFA4 Receptors
 CC NUMBER OF SEQUENCES: 6
 CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Genentech, Inc.
 CC STREET: 460 Point San Bruno Blvd
 CC CITY: South San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94080
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WinPatIn (Genentech)
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/076,093A
 CC FILING DATE: 11-Jun-1993
 CC CLASSIFICATION: 530
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/810782
 CC FILING DATE: 19-DEC-1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/677211
 CC FILING DATE: 29-MAR-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Love, Richard B.
 CC REGISTRATION NUMBER: 34,659
 CC REFERENCE/DOCKET NUMBER: 706P2
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415/225-5530
 CC TELEFAX: 415/952-9881
 CC TELEX: 910/371-7168
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 350 amino acids
 CC TYPE: Amino Acid
 CC TOPOLOGY: Linear
 CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 1; Length 350;
 Best Local Similarity 44.4%; Pred. No. 8.26e-79;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMELETLNKKVYIIAYALVLLSLGNSLVMLVILYRGVRSVTDVYLLNLALAD 85
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNILVITFAFYKARSDTDVYLLNMAID 83
 Db 86 LLFALTLPWAASKVNG-WIFGTFCKVSVLKEVNFYSGILLACISVDRIYLAIVHATR 144
 QY 84 ILFVLTLPFWAVSHATGAWFSNATCKLLGIYAINFCNCGMLLTCTISMDRYIAIVQATK 143
 Db 145 TLTKQ-RHL--VKFVLCGCGWLSMNLSPFLFRQAYHPNNSPVC---YEVIGNDTAKW 198
 QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
 Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTTLRTLFKAHMGQKHRAVRVFAVVLFLCWLPL 258
 QY 202 KLLMLGELLGFFPLMFIMFCYFIVKTLVQAQNSKRKRAIRVIAVVLFLACQIPH 261
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCLNPIIYAFIQGNFRHGFLK 318
 QY 262 NWLLVTAANLGMKMR-SCQSEKLGITKTVTEVLAFLHCLLPVLYAFIQGNFRNYFLK 320
 Db 319 IL 320
 QY 321 IL 322

RESULT 5
 ID US-08-202-056-1 STANDARD; PRT; 350 AA.
 XX
 AC xxxxxx
 CC
 CC

Sequence 1, Application US/08202056

[illegible]

RESULT	6
ID	US-08-410-453A-1 STANDARD; PRT; 350 AA.
XX	xxxxxx
AC	
DT	
DE	Sequence 1, Application US/08410453A
CC	Sequence 1, Application US/08410453A
CC	Patent No. 5767063
CC	GENERAL INFORMATION:
CC	APPLICANT: Lee, James,
CC	APPLICANT: Holmes, William E.,
CC	APPLICANT: Woods, William I.,
CC	TITLE OF INVENTION: Human PF4A Receptors and Their Use
CC	NUMBER OF SEQUENCES: 2
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Genentech, Inc.
CC	STREET: 1 DNA Way
CC	CITY: South San Francisco
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 94080
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: WinPatin (Genentech)
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/410,453A
CC	FILING DATE: 24-Mar-1995
CC	CLASSIFICATION: 514
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: 08/234494
CC	FILING DATE: 28-APR-1994
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: 07/62711
CC	FILING DATE: 29-MAR-1991
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Love, Richard B.
CC	REGISTRATION NUMBER: 34,659
CC	REFERENCE/DOCKET NUMBER: P0706C1D1
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: 650/225-5530
CC	TELEFAX: 650/953-9881
CC	INFORMATION FOR SEQ ID NO: 1:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 350 amino acids
CC	TYPE: Amino Acid
CC	TOPOLOGY: Linear
SQ	SEQUENCE 350 AA; 39805 MW; 660082 CN;
	Query Match 35.4%; Score 984; DB 1; Length 350;
	Best Local Similarity 44.4%; Pred. No. 8.26e-79;
	Matches 134; Conservative 76; Mismatches 80; Indels 1
Dd	27 Y-SPCMLETFLNKVVIIAYALVFLLSLGNLSVMVLVILSVGRSVDYVLLLA
Qy	25 YCAPCRSSGSPG-YLYIRIAYSLICVLGLNLIVITVFAYFKKARSMTDYVLL
Dd	86 LLFALTLPWAASKVNG-WIFGTFLCKVWSLLKEVNFYSGITLLACISVDRYLA
Qy	84 ILFLVTLPFWAVSHATGAWVSFNATCKLLKGIIYNFCNMGMLLTCTISMDRYIA
Dd	145 TLTK-RHL--VKFYCLGCWGLSMNLSLPPFLFRQAYPNNSSPVC---EYVLGH
Qy	144 SFRLRSRTLPRSKIIICLVWGSLVSISSSTFEVNKYNTQGS-D-VCEPKYQTVS
Dd	199 RMVLRIPLHTTGFIVPLFMVLCFGFTLRTLFKAHMGOKHRAMRVIPAWVLFL
Qy	202 KLLMGLLELGGFTPLMFMIFFVIKTVLQAQNSRRHRAIRVIIIAVLVFLVA

Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCNLPYIAFYQGNFRHGFLEK 318
QY 262 NMVLLTAANLGMNR-SCQSEKLGITKVTVEVLAFLHCLNPLVLAIFYQKFRNYELK 320
Db 319 IL 320
QY 321 IL 322
RESULT 7
ID US-08-450-393A-7 STANDARD; PRT; 350 AA.
XX
AC xxxxxx
DT
XX
XX
DE Sequence 7, Application US/08450393A
XX
XX Sequence 7, Application US/08450393A
CC Patent No. 5707815
CC
CC GENERAL INFORMATION:
CC APPLICANT: Charo, Israel
CC APPLICANT: Coughlin, Shaun
CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMAOTTRACTANT
CC TITLE OF INVENTION: PROTEIN RECEPTORS
CC
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
CC STREET: 5 Palo Alto Square
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94306-2155
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/450,393A
CC FILING DATE: May 25, 1995
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Cseri, Luann
CC REGISTRATION NUMBER: 31,822
CC REFERENCE/DOCKET NUMBER: UCAL-237/02US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-843-5165
CC TELEFAX: 415-8857-0663
CC TELEX: 380816COOLEYPA
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;
Query Match 35.4%; Score 984; DB 1; Length 350;
Best Local Similarity 44.4%; Pred. No. 8.26e-79;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
Db 27 Y-SPCMLETINKYVYIAVALFLLSLGNSLVMLVILYSRVGRSVTDVYLLNALAD 85
QY 25 YCAPCRSSGSPG-YLYRIAYSILCVLLGNILVITFAFYKRSNTDYLINMAIAD 83
Db 86 LLFALPLTWAASKVNG-WLGFIFLCKVSVLLKKEVNFYSGLILLACISVDYRLAIVHATR 144
QY 84 ILFVLPLPWAWSHAGAWFNSNATCKLLGVIYAINFCGMLLTCTISMDRIYIAIVQATK 143
Db 145 TLTKQK-RHL--VRFVCLGCGWGLSMNLSLPFFFLFRQAYHPNNSPVC---YEVLGNDTAKW 198

QY 144 SFPLRSRTLPKRIICLVVWGLSVIISSTFVFNQYNTQSD-VCEPKYQTVS-EPIRW 201
Db 199 RMVLRILPHTFGFTVPLFVMLFCYGETLTLFKAHMGQKHRAMRVIAVVLIFLCLWLPY 258
QY 202 KLMLGLELLFGFFIPLFMIFCYTFIVKTLVQAQSKRKRKAIRVIAVVLFLACQIPH 261
Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCNLPYIAFYQGNFRHGFLEK 318
QY 262 NMVLLTAANLGMNR-SCQSEKLGITKVTVEVLAFLHCLNPLVLAIFYQKFRNYELK 320
Db 319 IL 320
QY 321 IL 322
RESULT 8
ID PCT-US93-11153-19 STANDARD; PRT; 358 AA.
XX
AC xxxxxx
DT
XX
XX
DE Sequence 19, Application PC/TUS9311153
XX
XX Sequence 19, Application PC/TUS9311153
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 84
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11153
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 19:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 358 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 358 AA; 40690 MW; 701130 CN;
Query Match 34.6%; Score 961; DB 2; Length 358;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;
Db 40 FL-PTMYSICFVGLGNGLVVLTVYFPRKRLTMTDTLLNLAVADILFLLLPFWAYS- 97

```

      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
38 YLYRIAYSLICVLGGLGNLVITFAFYKKARSDMTDYLNNMAIADILFVLTPFWAVSH 97
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
98 AAKSWFVGHFCKLIFAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHARVLLISKL 157
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
98 ATGAWFSNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSFRLSRTLPRSKI 157
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
158 SCVGWILATVLSIPPELLYSID-LQRSSEQ-AMRCSLITEHVE-AFTTIOVAQMWIGFLV 214
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
158 ICLVWGLSVIISSTFVNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
215 PLLAMSFYLVIIITLQARNFERNKAIVIAVVVFIVFOLPYNGVVLAAOTVANFNIT 274
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
217 PLFMFICYFTFIVTLVQAQNSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGKM 275
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
275 SSTCELSKQLNIADYVTSYSLACVRCVNPFLYAFYIGVFRNDLFKFDLGLCL 327
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
276 NRSQSEKLGITKTVTEVLAFLHCLLPVLYAFYIGQFRNFKILKDLWCV 328

RESULT 9
ID US-08-153-848-19 STANDARD; PRT; 358 AA.
XX
AC xxxxxx
XX
DT
DE
Sequence 19, Application US/08153848
XX
CC Sequence 19, Application US/08153848
CC Patent No. 5759804
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,848
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5759804and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 19:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 358 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 358 AA; 40690 MW; 701130 CN;
```

```

Query Match 34.68; Score 961; DB 1; Length 358;
Best Local Similarity 45.78; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
40 FL-PIMYSIICFVGLLGNLVITFAFYKKARSDMTDYLNNMAIADILFVLTPFWAVSH 97
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
38 YLYRIAYSLICVLGGLGNLVITFAFYKKARSDMTDYLNNMAIADILFVLTPFWAVSH 97
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
98 AAKSWFVGHFCKLIFAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHARVLLISKL 157
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
98 ATGAWFSNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATKSFRLSRTLPRSKI 157
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
158 SCVGWILATVLSIPPELLYSID-LQRSSEQ-AMRCSLITEHVE-AFTTIOVAQMWIGFLV 214
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
158 ICLVWGLSVIISSTFVNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
215 PLLAMSFYLVIIITLQARNFERNKAIVIAVVVFIVFOLPYNGVVLAAOTVANFNIT 274
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
217 PLFMFICYFTFIVTLVQAQNSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGKM 275
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
275 SSTCELSKQLNIADYVTSYSLACVRCVNPFLYAFYIGVFRNDLFKFDLGLCL 327
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
276 NRSQSEKLGITKTVTEVLAFLHCLLPVLYAFYIGQFRNFKILKDLWCV 328

RESULT 10
ID PCT-US93-11153-15 STANDARD; PRT; 378 AA.
XX
AC xxxxxx
XX
DT
DE
Sequence 15, Application PC/TUS9311153
XX
CC Sequence 15, Application PC/TUS9311153
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11153
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 378 amino acids
CC TYPE: amino acid
```

CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 378 AA; 42873 MW; 782819 CN;

Query Match 34.6%; Score 961; DB 2; Length 378;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 60 FL-PIMYSIIICFVGLGNGLVLYIYFKRLKMTDTYLLNLAVALDILFLTLFPWAYS- 117
QY 38 YLYRIAYSLICVLGNGLVLYIYFAFKARSMYDVLLNMAIADILFVLTLPFWAVSH 97

Db 118 AAKSWFVGHFCKLIFAIYKMSFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 177
QY 98 ATGAWFVSNAATCKLKGIAINFCGMLLLCISMDRYIAIVQAKSFRLSRTLPRSKI 157

Db 178 SCVGIWILATVLSIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFITIQAQMVIGFLY 234
QY 158 ICLVWGLSVIISSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216

Db 235 PLLAMSCYLVIIIRTLQARNFERNKAIVIAVVVFIQVLPYNGVVLQAVANFNIT 294
QY 217 PLMFIFCYTFIVKTLVQAQSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGKM 275

Db 295 SSTCELSKQNLAIADVTYSLACVRCVNPFLYAFIGVFRNDLFKFLKDLGCL 347
QY 276 NRSQSEKLGITKTVTEVLAFLHCCLPVLYAFIGQFRNFYFLKILKDLWCVC 328

RESULT 11
ID US-08-153-848-15 STANDARD; PRT; 378 AA.
XX AC xxxxxx
DT
XX
DE Sequence 15, Application US/08153848
XX
CC Sequence 15, Application US/08153848
CC Patent No. 5759804
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,848
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5759804and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448

CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 378 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 378 AA; 42873 MW; 782819 CN;

Query Match 34.6%; Score 961; DB 1; Length 378;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 60 FL-PIMYSIIICFVGLGNGLVLYIYFKRLKMTDTYLLNLAVALDILFLTLFPWAYS- 117
QY 38 YLYRIAYSLICVLGNGLVLYIYFAFKARSMYDVLLNMAIADILFVLTLPFWAVSH 97

Db 118 AAKSWFVGHFCKLIFAIYKMSFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 177
QY 98 ATGAWFVSNAATCKLKGIAINFCGMLLLCISMDRYIAIVQAKSFRLSRTLPRSKI 157

Db 178 SCVGIWILATVLSIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFITIQAQMVIGFLY 234
QY 158 ICLVWGLSVIISSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216

Db 235 PLLAMSCYLVIIIRTLQARNFERNKAIVIAVVVFIQVLPYNGVVLQAVANFNIT 294
QY 217 PLMFIFCYTFIVKTLVQAQSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGKM 275

Db 295 SSTCELSKQNLAIADVTYSLACVRCVNPFLYAFIGVFRNDLFKFLKDLGCL 347
QY 276 NRSQSEKLGITKTVTEVLAFLHCCLPVLYAFIGQFRNFYFLKILKDLWCVC 328

RESULT 12
ID US-08-153-848-7 STANDARD; PRT; 410 AA.
XX AC xxxxxx
DT
XX
DE Sequence 7, Application US/08153848
XX
CC Sequence 7, Application US/08153848
CC Patent No. 5759804
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,848
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5759804and, Greta E.

CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC NAME: Noland, Greta E.
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 410 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 410 AA; 46569 MW; 921741 CN;

Query Match 34.6%; Score 961; DB 1; Length 410;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 92 EL-PMYSIIICFVGLLGNGLVLTYYFKRLKMTDTYLLNLAVADILFLTLPLFWAYS- 149
QY 38 YLRYAISLCVGLLGNLIVITFAFYKRSMTDVLNMAINDILFLTLPLFWAVSH 97
Db 150 AAKSWVGFHCFKLPFAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 209
QY 98 ATGANWFSNATCKLLKGIYAINFCGMLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157
Db 210 SCVGWILATVLSIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFTIIQVQAVMVGFLV 266
QY 158 ICLVWGLSVIISSTFVFNQKNTQSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
Db 267 PLLAMSFCLVLIIRTLQARNFERNKAIRVIIVAVVVFVIFQLPYNGVLAQTVANENIT 326
QY 217 PLMFMFICFTFIVKTLVQAQNSKRHRKAIRVIIVAVVFLACQIPHNNVLLV-TAANLGM 275
Db 327 SSTCELSKOLNIAYDVTYSLACVRCVNPFLYAFIGVKFRNDLFKFLKDLGCL 379
QY 276 NRSCQSEKLGITYTKTVTEVLAFLHCCNLPVLYAFIGKFRNFKILKDLWCV 328

RESULT 13
ID PCT-US93-11153-7 STANDARD; PRF: 410 AA.
XX XXXXX
XX

Sequence 7, Application PC/TUS9311153

CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald

CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.

CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive

CC CITY: Chicago
CC STATE: Illinois

CC COUNTRY: USA
CC ZIP: 60606

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11153

CC FILING DATE:
CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:
CC

CC APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.

CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300

CC TELEFAX: (312) 474-0448
CC TELEX: 25-3856

CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 410 amino acids

CC TYPE: amino acid
CC TOPOLOGY: linear

CC MOLECULE TYPE: protein
CC SEQUENCE 410 AA; 46569 MW; 921741 CN;

Query Match 34.6%; Score 961; DB 2; Length 410;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 92 EL-PMYSIIICFVGLLGNGLVLTYYFKRLKMTDTYLLNLAVADILFLTLPLFWAYS- 149
QY 38 YLRYAISLCVGLLGNLIVITFAFYKRSMTDVLNMAINDILFLTLPLFWAVSH 97
Db 150 AAKSWVGFHCFKLPFAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 209
QY 98 ATGANWFSNATCKLLKGIYAINFCGMLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157
Db 210 SCVGWILATVLSIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFTIIQVQAVMVGFLV 266
QY 158 ICLVWGLSVIISSTFVFNQKNTQSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
Db 267 PLLAMSFCLVLIIRTLQARNFERNKAIRVIIVAVVVFVIFQLPYNGVLAQTVANENIT 326
QY 217 PLMFMFICFTFIVKTLVQAQNSKRHRKAIRVIIVAVVFLACQIPHNNVLLV-TAANLGM 275
Db 327 SSTCELSKOLNIAYDVTYSLACVRCVNPFLYAFIGVKFRNDLFKFLKDLGCL 379
QY 276 NRSCQSEKLGITYTKTVTEVLAFLHCCNLPVLYAFIGKFRNFKILKDLWCV 328

RESULT 14
ID PCT-US95-00476-8 STANDARD; PRF: 355 AA.
XX XXXXX
XX

Sequence 8, Application PC/TUS9500476

CC GENERAL INFORMATION:
CC APPLICANT: The Regents of the University of California

CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMAATTRACTANT
CC NUMBER OF SEQUENCES: 14

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Robbins, Berliner & Carson

CC STREET: 201 N. Figueroa Street, 5th Floor
CC CITY: Los Angeles

CC STATE: California
CC COUNTRY: USA

CC ZIP: 90012-2628
CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/00476
CC FILING DATE:

CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berliner, Robert
CC REGISTRATION NUMBER: 20,121
CC REFERENCE/DOCKET NUMBER: 5555-291
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 310-977-1001
CC TELEFAX: 310-977-1003
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 355 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC HYPOTHETICAL: NO
CC SEQUENCE 355 AA; 40122 MW; 676446 CN;

Query Match 34.2%; Score 951; DB 2; Length 355;
Best Local Similarity 43.8%; Pred. No. 9,77e-76;
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;
Db 44 YFVVIYALVFLLSLGNLSVLMVILYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 103
QY 38 YLYRIAYSLICVLGLGNILVITFAFYKARSMTDVYLLNMAIADILFVLTLPFWAVSH 97
Db 104 VNG-WIFGFLCKVSVLLKEVNFYSIGILLACISVDRIYLAIVHATRLTQK-RYL--VKF 159
QY 98 ATGAWVFSNATCKLLKGIYAINFNGMLLTICISMDRYIAIVQATKSFRLSRTPRSKI 157
Db 160 ICLSWGLSLLALPVLFRRTVSSNVSPAC---YEDMGNNTANWMLRLILPQSGFI 216
QY 158 ICLVWGLSVIISSTFVFNQK-YNTQGDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215
Db 217 VPLLMLFCYGTLETLEKHAHQKRAVRVFAVVLIFLLCWLPYNLVLLADTLMTQV 276
QY 216 IPLMFIFCYTIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNKVLVYTAANLGM 275
Db 277 IQETCERNHIDRALDATEILGILHSCNPLIYAFIQGKFRHGLLKIL 324
QY 276 NR-SCQSEKLGITKVTVEVLAFLHCLNPNVLYAFIQGKFRNFKIL 322

RESULT 15
ID US-07-759-568-1 STANDARD; PRT; 355 AA.
XX
AC xxxxxx
XX

Sequence 1, Application US/07759568
Patent No. 5374506
GENERAL INFORMATION:
APPLICANT: Murphy, Philip M.
TITLE OF INVENTION: Cloning of cDNA Encoding a Functional
TITLE OF INVENTION: Human Interleukin-8 Receptor
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07759,568

CC FILING DATE: 19910913
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Scott, Watson T.
CC REGISTRATION NUMBER: 26581
CC REFERENCE/DOCKET NUMBER: WTS/5683/91535/WBH
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-861-3000
CC TELEFAX: 202-822-0944
CC TELEX: 6714627 cush
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 355 amino acids
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 355 AA; 40122 MW; 676446 CN;
Query Match 34.2%; Score 951; DB 1; Length 355;
Best Local Similarity 43.8%; Pred. No. 9,77e-76;
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;
Db 44 YFVVIYALVFLLSLGNLSVLMVILYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 103
QY 38 YLYRIAYSLICVLGLGNILVITFAFYKARSMTDVYLLNMAIADILFVLTLPFWAVSH 97
Db 104 VNG-WIFGFLCKVSVLLKEVNFYSIGILLACISVDRIYLAIVHATRLTQK-RYL--VKF 159
QY 98 ATGAWVFSNATCKLLKGIYAINFNGMLLTICISMDRYIAIVQATKSFRLSRTPRSKI 157
Db 160 ICLSWGLSLLALPVLFRRTVSSNVSPAC---YEDMGNNTANWMLRLILPQSGFI 216
QY 158 ICLVWGLSVIISSTFVFNQK-YNTQGDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215
Db 217 VPLLMLFCYGTLETLEKHAHQKRAVRVFAVVLIFLLCWLPYNLVLLADTLMTQV 276
QY 216 IPLMFIFCYTIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHNKVLVYTAANLGM 275
Db 277 IQETCERNHIDRALDATEILGILHSCNPLIYAFIQGKFRHGLLKIL 324
QY 276 NR-SCQSEKLGITKVTVEVLAFLHCLNPNVLYAFIQGKFRNFKIL 322

Search completed: Fri Nov 13 12:08:42 1998
Job time : 25 secs.

CC MOLECULE TYPE: protein
SQ SEQUENCE 365 AA; 41375 MW; 721306 CN;

Query Match 100.0%; Score 2779; DB 10; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.60e-231;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MFSTPVKIIICQSIHLITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60
|||||
Qy 1 MFSTPVKIIICQSIHLITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60
|||||
Db 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120
|||||
Qy 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120
|||||
Db 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180
|||||
Qy 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180
|||||
Db 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFFIPLMFMIFCYTFIVKTLVQAQNSKR 240
|||||
Qy 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFFIPLMFMIFCYTFIVKTLVQAQNSKR 240
|||||
Db 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITYTKTVEVLAFHLC 300
|||||
Qy 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITYTKTVEVLAFHLC 300
|||||
Db 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKYKSGFSCAGRYSENISROTSETADNDNA 360
|||||
Qy 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKYKSGFSCAGRYSENISROTSETADNDNA 360
|||||
Db 361 SSFTM 365
|||||
Qy 361 SSFTM 365
|||||

RESULT 2
ID US-08-887-977-10 STANDARD; PRT; 365 AA.
XX
AC xxxxxx
DT
XX
DE
CC Sequence 10, Application US/08887977
CC GENERAL INFORMATION:
CC APPLICANT: Wang, Wei
CC APPLICANT: Gish, Kurt C.
CC APPLICANT: Schall, Thomas J.
CC APPLICANT: Vicari, Alain P.
CC APPLICANT: Zlotnik, Albert
CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DNAX Research Institute
CC STREET: 901 California Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304-1104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/887,977
CC FILING DATE: 03-JUL-1997
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/021,644
CC FILING DATE: 05-JUL-1996
CC

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/028,329
CC FILING DATE: 11-OCT-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ching, Edwin P.
CC REGISTRATION NUMBER: 34,090
CC REFERENCE/DOCKET NUMBER: DX0589K1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650-852-9192
CC TELEFAX: 650-496-1200
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 365 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 365 AA; 41375 MW; 721306 CN;

Query Match 100.0%; Score 2779; DB 12; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.60e-231;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MFSTPVKIIICQSIHLITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60
|||||
Qy 1 MFSTPVKIIICQSIHLITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60
|||||
Db 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120
|||||
Qy 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120
|||||
Db 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180
|||||
Qy 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180
|||||
Db 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFFIPLMFMIFCYTFIVKTLVQAQNSKR 240
|||||
Qy 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFFIPLMFMIFCYTFIVKTLVQAQNSKR 240
|||||
Db 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITYTKTVEVLAFHLC 300
|||||
Qy 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITYTKTVEVLAFHLC 300
|||||
Db 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKYKSGFSCAGRYSENISROTSETADNDNA 360
|||||
Qy 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKYKSGFSCAGRYSENISROTSETADNDNA 360
|||||
Db 361 SSFTM 365
|||||
Qy 361 SSFTM 365
|||||

RESULT 3
ID US-09-045-583-48 STANDARD; PRT; 374 AA.
XX
AC xxxxxx
DT
XX
DE
CC Sequence 48, Application US/09045583
CC GENERAL INFORMATION:
CC APPLICANT: Graham, Gerard J. et al.
CC TITLE OF INVENTION: Novel Molecules of the G Protein-Coupled
CC NUMBER OF SEQUENCES: 56
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 28 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC

CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/045,583
CC FILING DATE: 20-MAR-98
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mandragoras, Amy E.
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: MNI-044
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)742-4214
CC INFORMATION FOR SEQ ID NO: 48:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 374 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 374 AA; 42494 MW; 759989 CN;

Query Match 88.0%; Score 2445; DB 14; Length 374;
Best Local Similarity 98.8%; Pred. No. 9,38e-202;
Matches 324; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
Db 47 LFVPTAYSLLICVGLLGNLTIVITFAFYKARSMTDYLNNMAIADILFLVTLFPWAVSH 106
Qy 39 LY-RIAYSLLICVGLLGNLTIVITFAFYKARSMTDYLNNMAIADILFLVTLFPWAVSH 97
Db 107 ATGAWFVSNTCKLLKGIXAINFNCGMLLTCTISMDRYIAIVQATKSFRLRSRTPRSKI 166
Qy 98 ATGAWFVSNTCKLLKGIXAINFNCGMLLTCTISMDRYIAIVQATKSFRLRSRTPRSKI 157
Db 167 ICLVWGLSVIISSTFVFNQKNTQGSVDCEPKYQTVSEPIRWKLLMLGLELLFGFFIP 226
Qy 158 ICLVWGLSVIISSTFVFNQKNTQGSVDCEPKYQTVSEPIRWKLLMLGLELLFGFFIP 217
Db 227 LMFIFCTFIIVKTLVQAKSRHKAIRVIAVVLFLACQIPIHNMVLLVTAANLGRNR 286
Qy 218 LMFIFCTFIIVKTLVQAKSRHKAIRVIAVVLFLACQIPIHNMVLLVTAANLGRNR 277
Db 287 SCQSEKLGITVTVEVLAFLHCLNPVLYAFIGQKFRNFTFLKILDLWCVRKRYKSSGF 346
Qy 278 SCQSEKLGITVTVEVLAFLHCLNPVLYAFIGQKFRNFTFLKILDLWCVRKRYKSSGF 337
Db 347 SCAGRYSENISQTSQTSADNDNASFTM 374
Qy 338 SCAGRYSENISQTSQTSADNDNASFTM 365

RESULT 4
ID US-08-802-627A-2 STANDARD; PRT; 350 AA.
XX xxxxxx
AC
DT
DE
XX
XX

Sequence 2, Application US/08802627A
Sequence 2, Application US/08802627A
GENERAL INFORMATION:
CC APPLICANT: Lee, James
CC APPLICANT: Wood, William I.
CC TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.

CC STREET: 460 Point San Bruno Blvd
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Winpatin (Genentech)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/802,627A
CC FILING DATE: 19-Feb-1997
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/284586
CC FILING DATE: 10-AUG-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/076093
CC FILING DATE: 11-JUN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/810782
CC FILING DATE: 19-DEC-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Love, Richard B.
CC REGISTRATION NUMBER: 34,659
CC REFERENCE/DOCKET NUMBER: P0706P2P1D2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415/225-5530
CC TELEFAX: 415/953-9881
CC TELEX: 910/371-7168
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: Amino Acid
CC TOPOLOGY: Linear
CC SEQUENCE 350 AA; 39805 MW; 560082 CN;

Query Match 35.4%; Score 984; DB 12; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
Db 27 Y-SPCMLETFILNKYVLIAYALVFLLSLGNLSVLMVILSVGRSVDYVLLNLAAD 85
Qy 25 YCAPCRSGSPG-YLYRIAYSLLICVGLLGNLTIVITFAFYKARSMTDYLNNMAIAD 83
Db 86 LFLALTPIWAASKVNG-WIFGTFLCKVVSLLKEVNSYSGILLACISVDYVLAIVHATR 144
Qy 84 ILFVLTLPFWAVSHATGAWFVSNTCKLLKGIXAINFNCGMLLTCTISMDRYIAIVQATK 143
Db 145 TLQK-RHL--VKFVCLGCGLSNMNLSLFFFLFRQAYHPNNSSPVC---YEVLGNDTAKW 198
Qy 144 SFLRSRTPRSKIICLVWGLSVIISSTFVFNQKNTQGS-DVCEPKYQTVS-EPFRW 201
Db 199 RMVRLIPHTFGFIVPLFVMLFCYGFRTLRTLFKAHMGOKHARMRVFAVFLFLCWLPY 258
Qy 202 KLLMLGLELLFGFFIPLFMIFCTFIIVKTLVQAKSRHKAIRVIAVVLFLACQIPH 261
Db 259 NLVLLADTLMETQVIOETCERRNNIGRALDATEILGFLHSLCLNPIIYAFIGNFRHGLK 318
Qy 262 NMVLLVTANLGRNR-SCQSEKLGITVTVEVLAFLHCLNPVLYAFIGQKFRNFTFLK 320
Db 319 IL 320
Qy 321 IL 322

RESULT 5
ID US-09-104-296-2 STANDARD; PRT; 350 AA.
XX xxxxxx
AC
DT

```
XX DE Sequence 2, Application US/09104296
XX CC Sequence 2, Application US/09104296
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Lee, James
CC CC APPLICANT: Wood, William I.
CC CC TITLE OF INVENTION: PF4A Receptors
CC CC NUMBER OF SEQUENCES: 6
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Genentech, Inc.
CC CC STREET: 1 DNA Way
CC CC CITY: South San Francisco
CC CC STATE: California
CC CC COUNTRY: USA
CC CC ZIP: 94080
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: WinPatIn (Genentech)
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/09/104,296
CC CC FILING DATE: 24-June-1998
CC CC CLASSIFICATION:
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/701265
CC CC FILING DATE: 22-AUG-1996
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/664228
CC CC FILING DATE: 06-JUN-1996
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/076093
CC CC FILING DATE: 11-JUN-1993
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 07/810782
CC CC FILING DATE: 19-DEC-1991
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Love, Richard B.
CC CC REGISTRATION NUMBER: 34,659
CC CC REFERENCE/DOCKET NUMBER: P0706P2C2
CC CC TELEPHONE: 415/225-5530
CC CC TELEFAX: 415/952-9881
CC CC TELEX: 910/371-7168
CC CC INFORMATION FOR SEQ ID NO: 2:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 350 amino acids
CC CC TYPE: Amino Acid
CC CC TOPOLOGY: Linear
CC CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 15; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSDVYLLNMAIAD 83
Db 86 LFLALTLPIWASKVNG-WIFGTFLCKVSVLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTLFPWAVSHATGAWVFSNATCKLLKGIYAINFNCGMILLTLCISMDRYIAIVQATK 143
Db 145 TLTKQ-RHL--VKFVCLGCGWLSNLSLPEFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLSRTPRSKIICLVVWGLSVIISSTFVNQKINTQGS-DVCEPKYQTVS-EPIRW 201
Db 199 RMYRLRPLHFGFTVPLFVLMFCYCGFTLRTLFKAHMGOKHRAMRVFAVWLFILCWLPPY 258
QY 202 KLLMLGLELFGFIPLFMFMFCYFTFVKTLVQAKNSKRHAIKRVIAVLVFLACQIPH 261
Db 259 NUVLADTLMTQVTOETCERRNNIGRALDATELGLFHLSCNLPYIYAFIGNGRHFLK 318

QY 262 NMVLLVTAANLGMNR-SCQSEKLGTYTFTVTEVLAFLHCLLNPLVLYAFIGKFRNYFLK 320
Db 319 IL 320
QY 321 IL 322

RESULT 6
ID US-08-450-393-7 STANDARD; PRT; 350 AA.
XX AC xxxxxx
XX CC
DT DT
DE DE
DE DE
XX CC Sequence 7, Application US/08450393
XX CC Sequence 7, Application US/08450393
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Charo, Israel
CC CC APPLICANT: Coughlin, Shaun
CC CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
CC CC TITLE OF INVENTION: PROTEIN RECEPTORS
CC CC NUMBER OF SEQUENCES: 14
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
CC CC STREET: 5 Palo Alto Square
CC CC CITY: Palo Alto
CC CC STATE: California
CC CC COUNTRY: USA
CC CC ZIP: 94308-2155
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/08/450,393
CC CC FILING DATE: May 25, 1995
CC CC CLASSIFICATION: 424
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Cseir, Luann
CC CC REGISTRATION NUMBER: 31,822
CC CC REFERENCE/DOCKET NUMBER: UCAL-237/02US
CC CC TELECOMMUNICATION INFORMATION:
CC CC TELEPHONE: 415-843-5165
CC CC TELEFAX: 415-8857-0663
CC CC TELEX: 380816CooleyPA
CC CC INFORMATION FOR SEQ ID NO: 7:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 350 amino acids
CC CC TYPE: amino acid
CC CC TOPOLOGY: linear
CC CC MOLECULE TYPE: protein
CC CC HYPOTHETICAL: NO
CC CC ANTI-SENSE: NO
CC CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 8; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSDVYLLNMAIAD 83
Db 86 LFLALTLPIWASKVNG-WIFGTFLCKVSVLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTLFPWAVSHATGAWVFSNATCKLLKGIYAINFNCGMILLTLCISMDRYIAIVQATK 143
Db 145 TLTKQ-RHL--VKFVCLGCGWLSNLSLPEFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLSRTPRSKIICLVVWGLSVIISSTFVNQKINTQGS-DVCEPKYQTVS-EPIRW 201
```


CC TELEFAX: 415/952-9881
CC TELEX: 910/371-7168
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: Amino Acid
CC TOPOLOGY: Linear
CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 6; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSVLMVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRRSGSPG-YLYRIAYSILICVLGLGNILVITFAFYKKARSMTDVYLLNMAIAD 83
Db 86 LLFALTLPWAASKVNG-WIFGTFCKVYSLKKEVNFYSGILLACISVDRIYAIHVATR 144
QY 84 ILFVLTLPFWAYSHATGANVFSNATCKLLGIYAINFCGMLLTCTISMDRYIAIVQATK 143
Db 145 TLTKQ-RHL--VKFVCLGCGWGLSMNLSLPPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQYNTQSD-VCEPKYQTVS-EPIRW 201
Db 199 RMVLRIILPHFTFGFIPLFVLMFCYGFTRTLFRAHMGOKHRAMRVIFAVVLIIFLLCWLPY 258
QY 202 KLLMLGLELFGFFIPLFMFICYTFIVKTLVQAQNSKRHKAIKRVIIAVVLFVLAQCIPIH 261
Db 259 NLVLADLTMRTOVIOETCERRNNIGRALDATEITGLFHLSCNLPYIAFYAGQFRHGFLK 318
QY 262 NMVLLVTANLGMNR-SCQSEKLIQYTKVTEVLAFLHCCNLPVIAFYAGQFRNYFLK 320
Db 319 IL 320
QY 321 IL 322

RESULT 9
ID PCT-US93-10672-3 STANDARD; PRT; 350 AA.
XX
AC xxxxxx
DT

Sequence 3, Application PC/TUS9310672

Sequence 3, Application PC/TUS9310672

GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
TITLE OF INVENTION: CC-CHEMOKINE RECEPTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10672
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.

CC REGISTRATION NUMBER: 35,136
CC REFERENCE/DOCKET NUMBER: 806
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415/225-1489
CC TELEFAX: 415/952-9881
CC TELEX: 910/371-7168

CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 1; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSVLMVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRRSGSPG-YLYRIAYSILICVLGLGNILVITFAFYKKARSMTDVYLLNMAIAD 83
Db 86 LLFALTLPWAASKVNG-WIFGTFCKVYSLKKEVNFYSGILLACISVDRIYAIHVATR 144
QY 84 ILFVLTLPFWAYSHATGANVFSNATCKLLGIYAINFCGMLLTCTISMDRYIAIVQATK 143
Db 145 TLTKQ-RHL--VKFVCLGCGWGLSMNLSLPPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQYNTQSD-VCEPKYQTVS-EPIRW 201
Db 199 RMVLRIILPHFTFGFIPLFVLMFCYGFTRTLFRAHMGOKHRAMRVIFAVVLIIFLLCWLPY 258
QY 202 KLLMLGLELFGFFIPLFMFICYTFIVKTLVQAQNSKRHKAIKRVIIAVVLFVLAQCIPIH 261
Db 259 NLVLADLTMRTOVIOETCERRNNIGRALDATEITGLFHLSCNLPYIAFYAGQFRHGFLK 318
QY 262 NMVLLVTANLGMNR-SCQSEKLIQYTKVTEVLAFLHCCNLPVIAFYAGQFRNYFLK 320
Db 319 IL 320
QY 321 IL 322

RESULT 10
ID US-08-805-478-2 STANDARD; PRT; 350 AA.
XX
AC xxxxxx
DT

Sequence 2, Application US/08805478

Sequence 2, Application US/08805478

GENERAL INFORMATION:
APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,478

CC FILING DATE: 25-FEB-1997
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/284586
CC FILING DATE: 10-AUG-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/076093
CC FILING DATE: 11-JUN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/810782
CC FILING DATE: 19-DEC-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Love, Richard B.
CC REGISTRATION NUMBER: 34,659
CC REFERENCE/DOCKET NUMBER: P07062P1C1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415/225-5530
CC TELEX: 910/371-7168
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: Amino Acid
CC TOPOLOGY: Linear
CC SEQUENCE 350 AA; 39805 MW; 660082 CN;
SQ
Query Match 35.4%; Score 984; DB 12; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
Db 27 Y-SPCMLETNLKYYVIAAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGLGNILVITFAFYKARSMTDYLNLMAIAD 83
Db 86 LFLALPLIWAASKVNG-WIFGTFCKVWSLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTPFWAVSHATGAWVFSNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATK 143
Db 145 TLTKR-RHL--VKFVCLGCGWLSMNLSPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTFLTLFKAHMGOKHRAMRVIFAVLIFLCLWLPY 258
QY 202 KLLMLGLELFGFFIPLMFIFCYTFIVKTLVQAQNSKRHKAIKRVIAVAVLFLACQIPH 261
Db 259 NLVLLADTLMTQVIOETCERNRNIGRALDATEILGFLHCLNPLIIYAFIGNFRHGFLEK 318
QY 262 NMVLLVTAANLGMKNR-SCQSEKLIGYTKTVEVLAFHCLNPLVYAFIGOKFRNYFLK 320
Db 319 IL 320
QY 321 IL 322
RESULT 11
ID US-08-410-456A-1 STANDARD; PRT: 350 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 1, Application US/08410456A
CC GENERAL INFORMATION:
CC APPLICANT: Lee, James,
CC APPLICANT: Holmes, William E.,
CC APPLICANT: Woods, William I.
CC TITLE OF INVENTION: Human PF4A Receptors and Their Use
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.

CC STREET: 1 DNA Way
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WinPatIn (Genentech)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/410,456A
CC FILING DATE: 24-Mar-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/234494
CC FILING DATE: 28-APR-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/677211
CC FILING DATE: 29-MAR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Love, Richard B.
CC REGISTRATION NUMBER: 34,659
CC REFERENCE/DOCKET NUMBER: P0706C1D2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650/225-5530
CC TELEX: 650/952-9881
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 350 amino acids
CC TYPE: Amino Acid
CC TOPOLOGY: Linear
CC SEQUENCE 350 AA; 39805 MW; 660082 CN;
SQ
Query Match 35.4%; Score 984; DB 8; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
Db 27 Y-SPCMLETNLKYYVIAAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGLGNILVITFAFYKARSMTDYLNLMAIAD 83
Db 86 LFLALPLIWAASKVNG-WIFGTFCKVWSLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTPFWAVSHATGAWVFSNATCKLLKGIYAINFNCGMLLTCTISMDRYIAIVQATK 143
Db 145 TLTKR-RHL--VKFVCLGCGWLSMNLSPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTFLTLFKAHMGOKHRAMRVIFAVLIFLCLWLPY 258
QY 202 KLLMLGLELFGFFIPLMFIFCYTFIVKTLVQAQNSKRHKAIKRVIAVAVLFLACQIPH 261
Db 259 NLVLLADTLMTQVIOETCERNRNIGRALDATEILGFLHCLNPLIIYAFIGNFRHGFLEK 318
QY 262 NMVLLVTAANLGMKNR-SCQSEKLIGYTKTVEVLAFHCLNPLVYAFIGOKFRNYFLK 320
Db 319 IL 320
QY 321 IL 322
RESULT 12
ID US-08-308-880-3 STANDARD; PRT: 350 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 3, Application US/08308880
XX
CC Sequence 3, Application US/08308880

CC GENERAL INFORMATION:
 CC APPLICANT: RICHARD HORUK, KULDEEP NEOTE, THOMAS SCHALL
 CC TITLE OF INVENTION: CC-CHEMOKINE RECEPTOR
 CC NUMBER OF SEQUENCES: 11
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Genentech, Inc.
 CC STREET: 460 Point San Bruno Blvd
 CC CITY: South San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94080
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: patin (Genentech)
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/308,880
 CC FILING DATE:
 CC CLASSIFICATION: 530
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/974025
 CC FILING DATE: 10-NOV-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Fitts, Renee A. 35,136
 CC REGISTRATION NUMBER:
 CC REFERENCE/DOCKET NUMBER: 806
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415/225-1489
 CC TELEFAX: 415/952-9881
 CC TELEX: 910/371-7168
 CC INFORMATION FOR SEQ ID NO: 3:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 350 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 7; Length 350;
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
 Db 27 Y-SPCMLETEFLNKVYVIAAYALVFLSLGNSLVMLVILYKSRGSRVTDVYLLNLALAD 85
 QY 25 YCAPRRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSDTDVYLLNMAIAD 83
 Db 86 LFLALTLPIWAASKVNG-WIFGTFLCKVSLKKEVNFYSGILLACISVDRIYAIHVATR 144
 QY 84 ILFVLTLFPWAVSHATGAWFVSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATK 143
 Db 145 TLTKQ-RHL--VKFVCLGCGWLSNLSLPPFLFRQAYHPNNSSPVC---YEVLGNDTAKW 198
 QY 144 SFRLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
 Db 199 RMVLRILPHTGFIPLFVLMFCYGTFTLTKFAHMGQKRAHMRVIFAVVLIPLCLWLPY 258
 QY 202 KLLMLGLELLEFGFFIPLMFIMFYCTFIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPH 261
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGLFHSCLNPIIYAFIQONFRHGFLK 318
 QY 262 NMVLLVTAANLGMKNR-SCQSEKLGITYTKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLK 320
 Db 319 IL 320
 QY 321 IL 322

RESULT 13
 ID US-08-441-562-3 STANDARD; PRF; 350 AA.
 XX
 AC
 XX
 XX
 DT

XX Sequence 3, Application US/08441562
 DE
 XX Sequence 3, Application US/08441562
 CC GENERAL INFORMATION:
 CC APPLICANT: RICHARD HORUK, KULDEEP NEOTE, THOMAS SCHALL
 CC TITLE OF INVENTION: Nucleic Acids Encoding C-C Chemokine Receptor
 CC TITLE OF INVENTION: and Vectors and Hosts Therefor
 CC NUMBER OF SEQUENCES: 11
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Genentech, Inc.
 CC STREET: 460 Point San Bruno Blvd
 CC CITY: South San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94080
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WinPatIn (Genentech)
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/441,562
 CC FILING DATE:
 CC CLASSIFICATION: 536
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/308880
 CC FILING DATE: 19-SEP-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/974025
 CC FILING DATE: 10-NOV-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Love, Richard B.
 CC REGISTRATION NUMBER: 34,659
 CC REFERENCE/DOCKET NUMBER: P0806C1D1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415/225-5530
 CC TELEX: 910/371-7168
 CC INFORMATION FOR SEQ ID NO: 3:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 350 amino acids
 CC TYPE: Amino Acid
 CC TOPOLOGY: Linear
 CC SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;
 Query Match 35.4%; Score 984; DB 8; Length 350;
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;
 Db 27 Y-SPCMLETEFLNKVYVIAAYALVFLSLGNSLVMLVILYKSRGSRVTDVYLLNLALAD 85
 QY 25 YCAPRRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSDTDVYLLNMAIAD 83
 Db 86 LFLALTLPIWAASKVNG-WIFGTFLCKVSLKKEVNFYSGILLACISVDRIYAIHVATR 144
 QY 84 ILFVLTLFPWAVSHATGAWFVSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATK 143
 Db 145 TLTKQ-RHL--VKFVCLGCGWLSNLSLPPFLFRQAYHPNNSSPVC---YEVLGNDTAKW 198
 QY 144 SFRLRSRTLPKRSKIICLVVWGLSVIISSTFVFNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
 Db 199 RMVLRILPHTGFIPLFVLMFCYGTFTLTKFAHMGQKRAHMRVIFAVVLIPLCLWLPY 258
 QY 202 KLLMLGLELLEFGFFIPLMFIMFYCTFIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPH 261
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGLFHSCLNPIIYAFIQONFRHGFLK 318
 QY 262 NMVLLVTAANLGMKNR-SCQSEKLGITYTKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLK 320
 Db 319 IL 320
 QY 321 IL 322


```

RESULT 14
ID US-08-801-238-2 STANDARD; PRT; 350 AA.
XX
AC
XX
XX
DT
XX
XX
DE
XX
Sequence 2, Application US/08801238
XX
Sequence 2, Application US/08801238
CC
GENERAL INFORMATION:
CC
APPLICANT: Lee, James
CC
TITLE OF INVENTION: PF4A RECEPTOR
CC
NUMBER OF SEQUENCES: 6
CC
CORRESPONDENCE ADDRESS:
CC
ADDRESSEE: Genentech, Inc.
CC
STREET: 460 Point San Bruno Blvd
CC
CITY: South San Francisco
CC
STATE: California
CC
COUNTRY: USA
CC
ZIP: 94080
CC
COMPUTER READABLE FORM:
CC
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC
COMPUTER: IBM PC compatible
CC
OPERATING SYSTEM: PC-DOS/MS-DOS
CC
SOFTWARE: WinPatIn (Genentech)
CC
CURRENT APPLICATION DATA:
CC
APPLICATION NUMBER: US/08/801,238
CC
FILING DATE: 19-Feb-1997
CC
CLASSIFICATION: 530
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 08/284586
CC
FILING DATE: 10-AUG-1994
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 08/076093
CC
FILING DATE: 11-JUN-1993
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 07/810782
CC
FILING DATE: 19-DEC-1991
CC
ATTORNEY/AGENT INFORMATION:
CC
NAME: Love, Richard B.
CC
REGISTRATION NUMBER: 34,659
CC
REFERENCE/DOCKET NUMBER: P0706P2P1D1
CC
TELECOMMUNICATION INFORMATION:
CC
TELEPHONE: 415/225-5530
CC
TELEFAX: 415/952-9881
CC
INFORMATION FOR SEQ ID NO: 2:
CC
SEQUENCE CHARACTERISTICS:
CC
LENGTH: 350 amino acids
CC
TYPE: Amino Acid
CC
TOPOLOGY: Linear
CC
SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 12; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETETLNKVVIIAYALVFLSLGNSLMVLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSGSSPG-YLYRIAYSLICVLGLGNILVWITFAFYKARSMTDYLNNWATAD 83
Db 86 LLFALTLPINWASKNG-WIFGTFLCKVYSLKLVNFVSGILLACISVDRIYLAIVHATR 144
QY 84 ILFVLTLPFWAVSHATGAWFVSNATCKLLKGIYAINFCGMLLLTICISMDRYAIVQATK 143
Db 145 TLTKQ-RHL--VKEVCLGCGCLSNLSLPFLFRQAYIPNNSPVC---YEVILGNDTAKW 198
QY 144 SFLRSRTLPRSKICLVVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EPIRW 201

Db 199 RMVLRILPHTFGFIYVPLFVMLFCYGYFTLRTLFKAHMGQKRAMRVLFVAVVLIFLCMLPY 258
QY 202 KLMLGLELLFGFFPLPMFMIFFYFIVATLVQAQSKRHKAIRVIAVVLVFLACQIPH 261
Db 259 NEVLLADTLMTQVVIQETCERRNNIRGALDATEILGLFHSCLNPIIYAFIQNFRHGFLK 318
QY 262 NMVLLVTAANLGMNR-SCQSEKLIGYTKTVTEVLAFLHCLLNPVLYAFIQGKFRNYFLK 320
Db 319 IL 320
QY 321 IL 322

RESULT 15
ID US-08-802-627-2 STANDARD; PRT; 350 AA.
XX
AC
XX
XX
DT
XX
XX
DE
XX
Sequence 2, Application US/08802627
CC
Sequence 2, Application US/08802627
CC
GENERAL INFORMATION:
CC
APPLICANT: Lee, James
CC
APPLICANT: Wood, William I.
CC
TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
CC
NUMBER OF SEQUENCES: 6
CC
CORRESPONDENCE ADDRESS:
CC
ADDRESSEE: Genentech, Inc.
CC
STREET: 460 Point San Bruno Blvd
CC
CITY: South San Francisco
CC
STATE: California
CC
COUNTRY: USA
CC
ZIP: 94080
CC
COMPUTER READABLE FORM:
CC
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC
COMPUTER: IBM PC compatible
CC
OPERATING SYSTEM: PC-DOS/MS-DOS
CC
SOFTWARE: WinPatIn (Genentech)
CC
CURRENT APPLICATION DATA:
CC
APPLICATION NUMBER: US/08/802,627
CC
FILING DATE: 19-Feb-1997
CC
CLASSIFICATION: 514
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 08/284586
CC
FILING DATE: 10-AUG-1994
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 08/076093
CC
FILING DATE: 11-JUN-1993
CC
PRIOR APPLICATION DATA:
CC
APPLICATION NUMBER: 07/810782
CC
FILING DATE: 19-DEC-1991
CC
ATTORNEY/AGENT INFORMATION:
CC
NAME: Love, Richard B.
CC
REGISTRATION NUMBER: 34,659
CC
REFERENCE/DOCKET NUMBER: P0706P2PID2
CC
TELECOMMUNICATION INFORMATION:
CC
TELEPHONE: 415/225-5530
CC
TELEFAX: 415/952-9881
CC
INFORMATION FOR SEQ ID NO: 2:
CC
SEQUENCE CHARACTERISTICS:
CC
LENGTH: 350 amino acids
CC
TYPE: Amino Acid
CC
TOPOLOGY: Linear
CC
SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 12; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETETLNKVVIIAYALVFLSLGNSLMVLVILYSRVGRSVTDVYLLNLALAD 85
```

QY	25	YCAPCRSSGSPG-YLYRIAYISLICVLGGLNLTWITFAFYKARSMTDVLNNAIAD	83
Db	86	LLFALTLPWAASKVNG-WIFGFTFLCKWSSLKEVNFYSGILLLACISVDORYLAIVHATR	144
QY	84	ILFVLTLFPWAVSHATCAWFSNATCKLLGYIAINFNCGMLLTTCISMDRYIAIQAAT	143
Db	145	TLTQK-RHL--VKFVCLGCGWGLSMNLSLPFFLFRQAYHPHNNSSPVC---YEVLGNDTAKW	198
QY	144	SFLRSRTLPSRKIIICLVWVGLSVIISSSTFVFNQKYNQTQSD-VCEPKYQTVS-EPINW	201
Db	199	RMVLRILPHPTFGFIVPLFVMLFCYGGTTLRTLFAHNGQKHARMVFAVWLFLCWLPL	258
QY	202	KLUMGLLELFGFFPLPMFIMFCYTFIVKTFLVQAQNSKRHKAIIRVIAVVLVFLACQIPH	261
Db	259	NVLVADTLMTORTVIOETGERRNNICRALDATEILGFLHSLNPIIYAFIGNFRHGLK	318
QY	262	NMVLVYTAANGKMN-RSCOSEKLIQYTKTVTEVLAFLHCLLNPIYAFIAGKFRNYFLK	320
Db	319	IL 320	
QY	321	IL 322	

Search completed: Fri Nov 13 12:11:13 1998
Job time : 132 secs.

!!SEQUENCE_LIST 1.0
(Nucleotide) WORDSEARCH of: /home/obryen/dra977/olig/us-08-887-977-9 check: 9278 from:
FROMIG of: /home/obryen/dra977/olig/US08887977.seq
sequence 9, application us/08887977
general information:
applicant: wang, wei
applicant: gish, kurt c.
applicant: schall, thomas j. . . .

TO: GenEMBL: * Sequences: 567,306 Total-length: 1,118,886,981 November 16, 1998 15:09

Database Release Information:

GenBank, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998
EMBL, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998

Word-size: 15 Words: 54754 Diagonals: 3,201 Total-diagonals: 2,000,000,000
Integral-width: 1 Alphabet: 4 List-size: 50 CPU minutes: 118.46

Sequence Strd Diag Score Width Documentation ..

GB_PR2:HSU45984 + 254 992 2 U45984 Homo sapiens CCR6 chemokine re
GB_PR1:HSCRL3 + 87 992 2 279784 H.sapiens G protein-coupled re
GB_PR3:HSU68032 + 222 990 2 U68032 Human G protein-coupled recept
GB_PR3:HSU68030 + 291 990 2 U68030 Human G protein-coupled recept
GB_PR1:HSU60000 + 27 930 2 U60000 Human IL8-related receptor (DR
GB_RO:AB009369 + 149 64 2 AB009369 Mus musculus mRNA for G prot
GB_IN:AC005148 + 894 19 1 AC005148 Drosophila melanogaster DNA
GB_IN:CEZK678 + 22866 19 1 279605 Caenorhabditis elegans cosmid
GB_HTG:AC004442 - 33655 19 1 AC004442 *** SEQUENCING IN PROGRESS *
GB_HTG:AC004311 - 29041 19 1 AC004311 *** SEQUENCING IN PROGRESS *
GB_RO:MMAJ3959 - 1325 19 1 AJ223959 Mus musculus mRNA for very-l
GB_RO:AF072760 - 1156 19 1 AF072760 Mus musculus fatty acid tran
GB_PAT:112527 - 1714 19 1 112527 Sequence 1 from patent US 5428
GB_PR2:AB013985 - -513 19 1 AB013985 Antirrhinum majus transposon
GB_STS:G39250 - -35 19 1 G39250 Z20386 zebrafish AB Daio re
GB_PR2:HS93H18 + 228 19 1 AF018631 Homo sapiens biotinidase (BT
GB_PR2:AC002542 - 69237 19 1 284488 Human DNA sequence from PAC 93
GB_HTG:HS347H13 - 136361 19 1 AC002542 Human BAC clone RG114A06 fi
GB_HTG:CEY7A5 + 116418 19 1 AL021576 Caenorhabditis elegans DNA
GB_IN:CFE26D10 + 17744 18 1 280223 Caenorhabditis elegans cosmid
GB_IN:CEC24H11 + 20715 18 1 281475 Caenorhabditis elegans cosmid
GB_IN:CEC24H11 + 24350 18 1 281475 Caenorhabditis elegans cosmid
GB_IN:AC004369 - 40361 18 1 AC004369 Drosophila melanogaster DNA
GB_IN:DMU63857 + 1152 18 1 U63857 Drosophila simulans decapenta
GB_IN:CET24A11 + 13516 18 1 249072 Caenorhabditis elegans cosmid
GB_IN:CELT20F5 + 27806 18 1 U80442 Caenorhabditis elegans cosmid
GB_IN:CEUHSF70 - 917 18 1 M18540 C.elegans heat shock protein 7
GB_IN:CEH31B20 + 9079 18 1 AL021066 Caenorhabditis elegans cosmi
GB_HTG:AC004556 - 92474 18 1 AC004556 *** SEQUENCING IN PROGRESS *
GB_HTG:AC004479 - 176462 18 1 AC004479 *** SEQUENCING IN PROGRESS
GB_VI:CVCP - 1647 18 1 X76716 Caliciviridae mRNA for capsid
GB_VI:CRNAORFS - 5321 18 1 X86557 Human enteric calicivirus ORF1
GB_RO:AB009369 + 152 18 1 AB009369 Mus musculus mRNA for G prot
GB_PR3:HSU85195 + 48083 18 1 U85195 Homo sapiens BAC129, complete
GB_PR3:HSB107C2 + 7446 18 1 AL008724 Human genomic DNA sequence f
GB_PR3:HS211A9 + 50916 18 1 295889 Human DNA sequence from clone
GB_PR3:AC004993 + 96033 18 1 AC004993 Homo sapiens PAC clone DJ119
GB_PL1:PYTRBCLH + 757 18 1 L34817 Pyramonas octopus chloroplas
GB_PL1:PYTRBCLG + 757 18 1 L34816 Pyramonas parkeae chloroplas
GB_PL1:PYTRBCLAD + 757 18 1 L34779 Pyramonas grossii chloroplas
GB_PL1:PYTRBCLAB + 757 18 1 L34777 Pyramonas propulsa chloropla
GB_PL1:PYTCPRBCLZ + 757 18 1 L34834 Pyramonas sp. chloroplast ri
GB_PL1:PEASHMTA + 1217 18 1 M87649 Pisum sativum serine hydroxyme
GB_PL1:ATU20624 + 674 18 1 U20624 Arabidopsis thaliana calcium-c
GB_BA1:U67522 + 5501 18 1 U67522 Methanococcus jannaschii secti
GB_PR2:HUAE000658 + 48083 18 1 AE000658 Homo sapiens T-cell receptor
GB_PR2:HS125H2 + 39027 18 1 298949 Human DNA sequence from clone
GB_PR2:AC002511 + 29886 18 1 AC002511 Human DNA from chromosome 19

of sequences containing motif is at least 15 nt long

UG-08-887-977-9. rgo

C 147	17	1.5	1339	22	MSL4GFBP1	M. musculus mRNA for in	5.90e+01
C 148	17	1.5	1503	23	MSL4GFBP1	MSL4GFBP1	5.90e+01
C 149	17	1.5	1526	17	R135P701	CDNA encoding R135P701	5.90e+01
C 150	17	1.5	1542	22	RATCFBP	Rat cytochrome P450 1a	5.90e+01
C 151	17	1.5	1550	20	HMGIC/GZEN	Human IgD and IgH germ	5.90e+01
C 152	17	1.5	1589	20	SHHG	Human gene for muscaric	5.90e+01
C 153	17	1.5	1590	23	SHHG	SHHG	5.90e+01
C 154	17	1.5	1590	23	RATCFBP2	Rat cytochrome P-450 b	5.90e+01
C 155	17	1.5	1747	16	AF000239	Xenopus laevis p1-act	5.90e+01
C 156	17	1.5	1986	21	SMPT7672	Macaca mulatta CC chem	5.90e+01
C 157	17	1.5	2001	20	HMPEL03	Human melanocyte prote	5.90e+01
C 158	17	1.5	2019	20	HMPEL03	Human melanocyte prote	5.90e+01
C 159	17	1.5	2049	23	HMPEL03	Human melanocyte prote	5.90e+01
C 160	17	1.5	2065	22	MUSHKBA	Mouse MHC class I H-2K	5.90e+01
C 161	17	1.5	2103	19	SCREK114X	C. cerevisiae REP114 ge	5.90e+01
C 162	17	1.5	2107	19	DD1ACTA03	Dicycotylellum discoid	5.90e+01
C 163	17	1.5	2109	20	DD1ACTA03	Dicycotylellum discoid	5.90e+01
C 164	17	1.5	2210	20	HUMACH28	Human gamma 2 inward	5.90e+01
C 165	17	1.5	2210	24	HUMMS1283	Human gamma 2 inward	5.90e+01
C 166	17	1.5	2337	22	RATCFBP45J	Rat cytochrome P450C9	5.90e+01
C 167	17	1.5	2411	16	CHTRPA	Chicken thymidine kina	5.90e+01
C 168	17	1.5	2411	16	CHTRPA	Chicken thymidine kina	5.90e+01
C 169	17	1.5	2543	20	SHHC7B1	H. sapiens immunoglob	5.90e+01
C 170	17	1.5	2543	20	SHHC7B3	H. sapiens immunoglob	5.90e+01
C 171	17	1.5	2543	20	SHHC7C	H. sapiens immunoglob	5.90e+01
C 172	17	1.5	2543	20	SHHC7C	H. sapiens immunoglob	5.90e+01
C 173	17	1.5	2543	20	SHHC7C	H. sapiens immunoglob	5.90e+01
C 174	17	1.5	2759	14	CEP60113	Caenorhabditis elegans	5.90e+01
C 175	17	1.5	2843	14	CEP60112	Caenorhabditis elegans	5.90e+01
C 176	17	1.5	3182	21	HUMHC91	Human Ig gamma1 J-mu	5.90e+01
C 177	17	1.5	3269	21	HUMHC91	Human mRNA for KIAA00	5.90e+01
C 178	17	1.5	3269	21	HUMHC91	Human mRNA for KIAA00	5.90e+01
C 179	17	1.5	3423	17	CEP60113	Caenorhabditis elegans	5.90e+01
C 180	17	1.5	3423	17	CEP60113	Caenorhabditis elegans	5.90e+01
C 181	17	1.5	3526	17	D76444	CDNA encoding R135P701	5.90e+01
C 182	17	1.5	4107	22	MM20238	Mus musculus GTFase p	5.90e+01
C 183	17	1.5	4107	22	MM20238	Mus musculus GTFase p	5.90e+01
C 184	17	1.5	5221	23	MUSHC64	Mouse MHC class I B-2	5.90e+01
C 185	17	1.5	5479	22	MUSHC64	Mouse MHC class I B-2	5.90e+01
C 186	17	1.5	5593	22	MUSHC64	Mouse MHC class I P5	5.90e+01
C 187	17	1.5	6036	21	SHHC7A1G	H. sapiens DNA for BCL7	5.90e+01
C 188	17	1.5	6036	21	SHHC7A1G	H. sapiens DNA for BCL7	5.90e+01
C 189	17	1.5	6036	21	SHHC7A1G	H. sapiens DNA for BCL7	5.90e+01
C 190	17	1.5	6581	16	SSGSH	S. salar gene for growth	5.90e+01
C 191	17	1.5	7252	20	HSNARY	Human mRNA for ankyrin	5.90e+01
C 192	17	1.5	7422	21	AF005962	Homo sapiens CC-chemok	5.90e+01
C 193	17	1.5	7422	21	AF005962	Homo sapiens CC-chemok	5.90e+01
C 194	17	1.5	9221	14	DMALD	D. melanogaster alpha-2	5.90e+01
C 195	17	1.5	11240	12	AE000864	Methanobacterium therm	5.90e+01
C 196	17	1.5	12663	12	ECBAE000168	Zscherichia coli K-12	5.90e+01
C 197	17	1.5	14360	1	ECBAE000604	Helicobacter pylori se	5.90e+01

18-08-887-977-9.198

[illegible]

351	1.4	995 27	PHL92G	Prospect Hill virus L	2.97e+02
352	1.4	1003 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
353	1.4	1004 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
354	1.4	1050 21	BOV1524	Bovine beta-casein mRNA	2.97e+02
355	1.4	1059 15	BOVCSBA	Bovine beta-casein mRNA	2.97e+02
356	1.4	1119 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
357	1.4	1120 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
358	1.4	1250 20	HMPMSN	Human cellular fas proto	2.97e+02
359	1.4	1259 27	ROTHICP	Adult diarrheal rotavi	2.97e+02
360	1.4	1285 17	MDM2478	Sequence 4 from Patent	2.97e+02
361	1.4	1286 22	MDM2478	Sequence 4 from Patent	2.97e+02
362	1.4	1286 22	MDM2478	Sequence 4 from Patent	2.97e+02
363	1.4	1334 16	GAVRX14	Gallus avris gene, ex	2.97e+02
364	1.4	1334 16	GAVRX14	Gallus avris gene, ex	2.97e+02
365	1.4	1335 16	GAVRX14	Gallus avris gene, ex	2.97e+02
366	1.4	1335 16	GAVRX14	Gallus avris gene, ex	2.97e+02
367	1.4	1337 16	SPASRGEN	Gallus avris gene, ex	2.97e+02
368	1.4	1337 16	SPASRGEN	Gallus avris gene, ex	2.97e+02
369	1.4	1349 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
370	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
371	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
372	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
373	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
374	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
375	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
376	1.4	1354 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
377	1.4	1441 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
378	1.4	1441 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
379	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
380	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
381	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
382	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
383	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
384	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
385	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
386	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
387	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
388	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
389	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
390	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
391	1.4	1454 12	STRAG173	S. pneumoniae (strain K	2.97e+02
392	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
393	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
394	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
395	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
396	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
397	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
398	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
399	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
400	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
401	1.4	1474 22	MDM2478	Mus musculus alpha-1-a	2.97e+02

453	1.4	2360 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
454	1.4	2360 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
455	1.4	2360 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
456	1.4	2360 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
457	1.4	2360 22	MDM2478	Mus musculus alpha-1-a	2.97e+02
458	1.4	2484 20	HOMLOX5	Human lipoproteinase mRNA	2.97e+02
459	1.4	2484 20	HOMLOX5	Human lipoproteinase mRNA	2.97e+02
460	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
461	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
462	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
463	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
464	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
465	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
466	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
467	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
468	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
469	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
470	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
471	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
472	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
473	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
474	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
475	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
476	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
477	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
478	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
479	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
480	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
481	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
482	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
483	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
484	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
485	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
486	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
487	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
488	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
489	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
490	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
491	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
492	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
493	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
494	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
495	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
496	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
497	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
498	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
499	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
500	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
501	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
502	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02
503	1.4	2500 17	MDM2478	Mus musculus alpha-1-a	2.97e+02

Tue Nov 17 08:55:24 1998

[illegible]

Tue Nov 17 08:55:24 1998

db	268	ACCTTTCGCTTTTAAAGAGGCGAGTCTATCAGACAGCTATATCTTGAACATGGCG	337
db	269	ACCTTTCGCTTTTAAAGAGGCGAGTCTATCAGACAGCTATATCTTGAACATGGCG	338
oy	181	ACCTTTCGCTTTTAAAGAGGCGAGTCTATCAGACAGCTATATCTTGAACATGGCG	240
db	328	ATTGAGACATCTCTTTTCTTACTCTCCCATTTTGGCAGTAGTACATCCATGGT	387
oy	441	ATTGAGACATCTCTTTTCTTACTCTCCCATTTTGGCAGTAGTACATCCATGGT	499
oy	329	ATTGAGACATCTCTTTTCTTACTCTCCCATTTTGGCAGTAGTACATCCATGGT	300
db	388	GGTGGGCTTTTCAGATATCCGACATCCGATCGATGTAAAGGACCTATATGCCATGCTT	447
oy	301	GGTGGGCTTTTCAGATATCCGACATCCGATGTAAAGGACCTATATGCCATGCTT	360
db	448	AATCGGCGAGATCTCTCTTCACTCTATCATATATGACAGCGGTATATCCCTATGACAG	507
oy	361	AATCGGCGAGATCTCTCTTCACTCTATCATATATGACAGCGGTATATCCCTATGACAG	420
db	508	GGCATATGATATCTCTGATCTGGACATCCGATCGGATGACATATATCCGATG	567
oy	421	GGCATATGATATCTCTGATCTGGACATCCGATCGGATGACATATATCCGATG	480
db	568	GTCTTGGGCGGCTGTATGATATATATATATATATATATATATATATATATATATAT	627
oy	481	GTCTTGGGCGGCTGTATGATATATATATATATATATATATATATATATATATATAT	540
db	624	ATACACGACAGGACAT	687
oy	541	ATACACGACAGGACAT	600
db	688	TGAAATCTGTATATCTTGGGCTGATGATCTATCTTGGCTGATATATATATATATATAT	747
oy	601	TGAAATCTGTATATCTTGGGCTGATGATCTATCTTGGCTGATATATATATATATATAT	660
db	748	ATGATATTTTGTACAGCTAT	807
oy	661	ATGATATTTTGTACAGCTAT	720
db	808	CAAAAGCATCTGCTGAT	867
oy	721	CAAAAGCATCTGCTGAT	780
db	868	CATACATCTGCTGCTTTTGAAGGCTGAAATTTGGGTAAATATGACCAATCTCTCGAG	927
oy	781	CATACATCTGCTGCTTTTGAAGGCTGAAATTTGGGTAAATATGACCAATCTCTCGAG	840
db	928	AGCAGAAAGCTATTTGGCTATATGAAATATCTATATATATATATATATATATATATAT	987
oy	841	AGCAGAAAGCTATTTGGCTATATGAAATATCTATATATATATATATATATATATATAT	900
db	988	TCTCTTAAAGCTCTGTCTATAGCTTTTATGGGACAGATGTACAGACTGTATATGACAG	1047
oy	901	TCTCTTAAAGCTCTGTCTATAGCTTTTATGGGACAGATGTACAGACTGTATATGACAG	960

Tue Nov 17 08:55:24 1998

[illegible]

Tue Nov 17 08:55:24 1998

QY	901	TCCTGACCTCTCTCTACCTCTTTATTCGGCAGATGTACGAACACTTCTTCGAG	960
Db	1048	ATTCTTAGGAACCTCTGTGTGTGAGAGAGATGACAGTCTCTGAGCTCTCTCTGCC	1107 *
QY	961	ATTCTTAGGAACCTCTGTGTGTGAGAGAGATGACAGTCTCTGAGCTCTCTCTGCC	1020
QY	1108	GCGAGTACTCTGACAAATCAATCTTCCTGGCAGACATGACGCGAGATGACGAGTGGG	1167
QY	1021	GCGAGTACTCTGACAAATCAATCTTCCTGGCAGACATGACGCGAGATGACGAGTGGG	1080
Db	1168	TGCTCTCTACATATCTGTATGAGAGCTGAGTCTCTCTAA	1206
QY	1081	TGCTCTCTACATATCTGTATGAGAGCTGAGTCTCTCTAA	1119
LOCUS	2		
DEFINITION	HS045984	3693 bp DNA	PRI
ACCESSION	U00948	Human sapiens CCR6 chemokine receptor (CCR6) gene, complete cds.	09-JUL-1997
VERSION	1		
KEYWORDS	Q2246432		
SOURCE	human.		
ORGANISM	human sapiens		
REFERENCE	1	Okuyama, T. Mitochondrial eukaryotes: Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Pages 1 to 3693.	
REFERENCE	2	Baba, I., Imai, K., Kishimura, M., Katagiri, M., Takagi, S., Hasehima, K., Homiyama, H. and Yoshie, O.	
AUTHORS	1	Identification of CCR6, the specific receptor for a novel lymphocyte-directed CC chemokine LARC	
JOURNAL	97131465	Chem. 272 (23), 14959-14968 (1997)	
MEDLINE	97131465		
REFERENCE	2	(bases 1 to 3693)	
AUTHORS	1	Latens, L. L., Modi, M. and Bonner, T. I.	
JOURNAL	Unpublished	Genomewide Cloning and Chromosomal Localization of a Potential G-Protein-Linked Receptor	
REFERENCE	3	(bases 1 to 3693)	
AUTHORS	1	Bonner, T. I.	
JOURNAL	Submitted (15-JAN-1996)	Tom I. Bonner, Lab of Cell Biology, NIH, Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA	
FEATURES	source	Location/Qualifiers	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="6"	
		/align="GPR-CY4"	
		/size(35..140,237..1136)	
	ERRA		

* This sequence is unfinished. When sequencing is complete, it will be added to the Genbank database.
* by a single finished sequence with the same accession number.

FEATURES
source
1. 90557
/organism="Caenorhabditis elegans"
/chromosome="X"
/clone="Y7A5"
/c1one="10455 g 19700 t 29702 others

BASE COUNT 20069 a 10631 c 10455 g 19700 t 29702 others

ORIGIN
Query Match 1.76; Score 19; DB 13; Length 90557;
Best Local Similarity 100.0%; Pred. No. 1.77e+00;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 83816 TTGACTGACAAATAT 83814
|||||
CP 40 TTGACTGACAAATAT 22

RESULT 15
LOCUS AC005148 95262 bp DNA INV 20-JUN-1998
DEFINITION Drosophila melanogaster DNA sequence (Pis D505759 (D204) and D509033 (D211))
ACCESSION AC003853 AC003854 AC003855 AC003856 AC003857
NID 93242681
KEYWORDS RTG
SOURCE Drosophila melanogaster (Subclones in tet from P1 clones D505759 D204 and D509033 (D211))
ORGANISM Drosophila melanogaster
REFERENCE
Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta: Pterygota: Diptera: Brachycera: Muscomorpha: Ephydroidea: Drosophila: 95262 bp
AUTHORS Celiker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blase, R.G., Chev.M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A., Humati, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B., Nixon, K.K., Pacle, B., Park, S., Pfeiffer, B., Punch, D., Shirk, E., Toney, B., Wan, K.R., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L. and Kimmel, B.
Sequencing of Drosophila chromosome 2L, region 50d11-50d22 (bases 1 to 95262)
JOURNAL
REFERENCE
AUTHORS Celiker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blase, R.G., Chev.M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A., Humati, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B., Nixon, K.K., Pacle, B., Park, S., Pfeiffer, B., Punch, D., Shirk, E., Toney, B., Wan, K.R., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L. and Kimmel, B.

Drosophilidae: Drosophila.
1 (bases 1 to 121479)
AUTHORS Celiker, S.E., George, R.A., Galle, E., Svirskas, R.R., Hoskins, R.A., Chev.M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A., Humati, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B., Lomtan, M.A., Mak, J., Maeda, P., Mok, M.S., Moshrefi, A.R., Nixon, K.K., Pacle, B., Park, S., Pfeiffer, B., Punch, D., Shirk, E., Toney, B., Wan, K.R., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L. and Kimmel, B.
Sequencing of Drosophila melanogaster
Unpublished
Cell lines: 121479
AUTHORS Aghavani, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blase, R.G., Chev.M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A., Humati, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B., Moshrefi, A.R., Mok, M.S., Moshrefi, A.R., Nixon, K.K., Pacle, B., Park, S., Pfeiffer, B., Punch, D., Shirk, E., Toney, B., Wan, K.R., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L. and Kimmel, B.
Direct Submission 1998
Submitted (MS 64-121) Berkeley, CA 94720, USA
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.hgc.lbl.gov/sequence-archive.html>) or submit a query to the HGC database. All contigs in this submission are complete. The following statistics are for the entire submission:
computed error rate <= 1/10.
*** WARNING: Phase 1 High Throughput Genome Sequence ***
* This sequence is unfinished. It consists of 33 contigs for which the order is not known. Their order in this record is arbitrary. In some cases, the exact lengths of the gaps between the contigs are also unknown; these gaps are presented as runs of N as a convenience. When sequencing is complete, the runs of N will be replaced by the actual sequence.
* by a single finished sequence with the same accession number.
* 1 622: contig of 622 bp in length
* 623 702: gap of unknown length
* 1965 2044: gap of unknown length
* 2045 2820: contig of 776 bp in length
* 2821 2900: gap of unknown length
* 2901 3552: contig of 652 bp in length
* 3553 3673: gap of unknown length
* 3673 4813: contig of 1141 bp in length
* 4814 4893: gap of unknown length
* 4894 5752: contig of 859 bp in length

Toney, B., Wan, K.R., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L. and Kimmel, B.
Submitted (MS 64-121) Berkeley, CA 94720, USA
Submitted (20-JUN-1998) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://fruitfly.berkeley.edu/sequence-archive.html>) or send email to drosophilahgc.lbl.gov.
Library locations: 138,60, 9, 95.
FEATURES
source
1. 95262
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="X"
/clone="Pis D505759 (D204) and D509033 (D211)"
/c1one="D505759 (D204) extends from bp 1 to bp 81035 and D509033 (D211) extends from bp 80404 to bp 95264"
BASE COUNT 28107 a 18978 c 19928 g 28149 t
ORIGIN

Query Match 1.76; Score 19; DB 14; Length 95262;
Best Local Similarity 100.0%; Pred. No. 1.77e+00;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1610 AAGGCGCAAGCCATCG 1628
|||||
CY 716 AAGGCGCAAGCCATCG 734

RESULT 16
LOCUS AC004442 121479 bp DNA RTG 21-JUL-1998
DEFINITION seq. SEQUENCING IN PROGRESS *** D504250; D48; Drosophila melanogaster chromosome 2L; Library location 45-36; RTGS phase 1, 33 unordered pieces
ACCESSION AC004442 AC003928 AC001356 AC001850 L81570 AC000870 AC002894 AC001357 AC002895 AC001851 L81569 AC000871 AC001329 AC000872 L81571 AC001852 AC000873 AC003618 AC000874 AC002896 AC001358
NID RTG: RTGS_PHASE1
KEYWORDS fruit fly
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
REFERENCE
Pterygota: Diptera: Brachycera: Muscomorpha: Ephydroidea.

5753 5832: gap of unknown length
* 5833 6744: contig of 912 bp in length
* 6745 7001: gap of unknown length
* 7002 7001: contig of 897 bp in length
* 7002 8937: contig of 1116 bp in length
* 8938 9017: gap of unknown length
* 9018 10013: contig of 996 bp in length
* 10014 11328: gap of unknown length
* 11329 11400: gap of unknown length
* 11401 12212: contig of 812 bp in length
* 12213 12292: gap of unknown length
* 12293 13763: gap of 1390 bp in length
* 13764 14950: contig of 1188 bp in length
* 14951 15030: gap of unknown length
* 15031 16666: contig of 1636 bp in length
* 16667 18546: contig of 1800 bp in length
* 18547 18628: gap of unknown length
* 18629 20030: contig of 1404 bp in length
* 20031 20110: gap of unknown length
* 20111 22226: gap of unknown length
* 22227 24562: contig of 2337 bp in length
* 24563 24642: gap of unknown length
* 24643 26781: contig of 1531 bp in length
* 26782 27701: contig of 1448 bp in length
* 27702 27781: gap of unknown length
* 27782 29546: contig of 1765 bp in length
* 29547 29626: gap of unknown length
* 29627 31103: gap of unknown length
* 31104 33020: contig of 1717 bp in length
* 33021 33100: gap of unknown length
* 33101 34911: contig of 1811 bp in length
* 34912 36973: contig of 1982 bp in length
* 36974 37053: gap of unknown length
* 37054 39149: contig of 2096 bp in length
* 39150 39228: gap of unknown length
* 39229 41065: gap of unknown length
* 41066 43398: contig of 2334 bp in length
* 43399 43400: gap of unknown length
* 43401 45566: contig of 2087 bp in length
* 45567 48937: gap of unknown length
* 48938 49857: contig of 4311 bp in length
* 49858 49857: gap of unknown length

COMMIT

SUBMITTED BY:
Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:sapiens@genome.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, and all ambiguities were resolved by resequencing. If a sequence from more than one subclone, and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
This clone was established as part of a mapping and sequencing collaboration between the NIGMS Chromosome 2 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information contact John D. McPherson at jdm@wustl.edu or Eric D. Green at egreen@nigms.nih.gov or send mail to: egreen@nigms.nih.gov

SOURCE INFORMATION:
This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 9789K. See: Shitaya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992), Kim et al., Genomics 14:213-8 (1996). The clone is a subclone of the research Genetics, Inc. (http://www.resgen.com).
VECTOR: pUC19

NEIGHBORING SEQUENCE INFORMATION:
This clone is adjacent to the right of B14A06. The actual start of this clone is at base position of B14A06. The orientation is unknown.

This clone contains STS's: aMS5377 (NID:g398836), aMS5845 (NID:g484312) and aMS5844 (NID:g191637).

Location/Qualifiers
1. 188741
/organism="Homo sapiens"
/db_xref="taxon:9606"

repeat_region 74413..74440
/note="7 copies of 4 mer 93 & conserved"
repeat_region 74932..75104
/note="MIR repeat: matches 18..201 of consensus"
repeat_region 75246..75549
/note="MIR repeat: matches 301..1 of consensus"
repeat_region 76493..76774
/note="MIR repeat: matches 1..1 of consensus"
repeat_region 78678..79111
/note="MIR repeat: matches 904..439 of consensus"
repeat_region 81131..81215
/note="MIR repeat: matches 1927..1843 of consensus"
repeat_region 81852..82274
/note="MIR repeat: matches 486..47 of consensus"
repeat_region 82321..82366
/note="23 copies of 2 mer 100 & conserved"
Note: remainder of annotations omitted.

Query Match
Query Local Similarity 100.0%; Score 19; DB 21; Length 160378;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69566 AGTTCCTAAGGACATCA 69584
QY 329 AGTTCCTAAGGACATCA 347

RESULT 18 AC002542 188741 bp DNA PRI 12-SEP-1997
LOCUS AC002542 188741 bp DNA PRI 12-SEP-1997
ACCESSION AC002542
NID AC002542
KEYWORDS HGC.

SOURCE
Human.
ORGANISM Human.
Eukaryote.
Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE
1 (bases 1 to 188741)
Davidson S, Rohlfing T, David M and Ahrens C.
JOURNAL Unpublished (1997) sapiens BAC clone R014A06
REFERENCE
2 (bases 1 to 188741)
Waterston R.
Direct Submission
TITLE Direct Submission
AUTHORS
JOURNAL
UNIVERSITY, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

repeat_region 74413..74440
/note="7 copies of 4 mer 93 & conserved"
repeat_region 74932..75104
/note="MIR repeat: matches 18..201 of consensus"
repeat_region 75246..75549
/note="MIR repeat: matches 301..1 of consensus"
repeat_region 76493..76774
/note="MIR repeat: matches 1..1 of consensus"
repeat_region 78678..79111
/note="MIR repeat: matches 904..439 of consensus"
repeat_region 81131..81215
/note="MIR repeat: matches 1927..1843 of consensus"
repeat_region 81852..82274
/note="MIR repeat: matches 486..47 of consensus"
repeat_region 82321..82366
/note="23 copies of 2 mer 100 & conserved"
Note: remainder of annotations omitted.

Query Match
Query Local Similarity 100.0%; Score 19; DB 21; Length 160378;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69566 AGTTCCTAAGGACATCA 69584
QY 329 AGTTCCTAAGGACATCA 347

RESULT 18 AC002542 188741 bp DNA PRI 12-SEP-1997
LOCUS AC002542 188741 bp DNA PRI 12-SEP-1997
ACCESSION AC002542
NID AC002542
KEYWORDS HGC.

SOURCE
Human.
ORGANISM Human.
Eukaryote.
Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE
1 (bases 1 to 188741)
Davidson S, Rohlfing T, David M and Ahrens C.
JOURNAL Unpublished (1997) sapiens BAC clone R014A06
REFERENCE
2 (bases 1 to 188741)
Waterston R.
Direct Submission
TITLE Direct Submission
AUTHORS
JOURNAL
UNIVERSITY, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

repeat_region 74413..74440
/note="7 copies of 4 mer 93 & conserved"
repeat_region 74932..75104
/note="MIR repeat: matches 18..201 of consensus"
repeat_region 75246..75549
/note="MIR repeat: matches 301..1 of consensus"
repeat_region 76493..76774
/note="MIR repeat: matches 1..1 of consensus"
repeat_region 78678..79111
/note="MIR repeat: matches 904..439 of consensus"
repeat_region 81131..81215
/note="MIR repeat: matches 1927..1843 of consensus"
repeat_region 81852..82274
/note="MIR repeat: matches 486..47 of consensus"
repeat_region 82321..82366
/note="23 copies of 2 mer 100 & conserved"
Note: remainder of annotations omitted.

Query Match
Query Local Similarity 100.0%; Score 19; DB 21; Length 160378;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69566 AGTTCCTAAGGACATCA 69584
QY 329 AGTTCCTAAGGACATCA 347

RESULT 18 AC002542 188741 bp DNA PRI 12-SEP-1997
LOCUS AC002542 188741 bp DNA PRI 12-SEP-1997
ACCESSION AC002542
NID AC002542
KEYWORDS HGC.

SOURCE
Human.
ORGANISM Human.
Eukaryote.
Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE
1 (bases 1 to 188741)
Davidson S, Rohlfing T, David M and Ahrens C.
JOURNAL Unpublished (1997) sapiens BAC clone R014A06
REFERENCE
2 (bases 1 to 188741)
Waterston R.
Direct Submission
TITLE Direct Submission
AUTHORS
JOURNAL
UNIVERSITY, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

[illegible][illegible]

```

BASE COUNT      1107 a 1143 c 1023 g 965 t
ORIGIN
Query Match      1.6%
Best Local Similarity 100.0%
Pred. No. 1.07+01
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Score 18; DB 14; Length 4238;

```

```

Query Match      1.69;      Score 18;  DB 14;  Length 4238;
Best Local Similarity 100.0%;  Pred. NO. 1.07e-01;
Matches 18;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
ORIGIN

```

gene
transcript="brain"
498..3782
/gene="KIA0316"
498..3782
/gene="KIA0316"
/cds_start=1
/cds_end=3782
/db_xref="db:022457"
translation="MTASGGRDYPVHNTQNPDPALSSCQIPPAKPKVYHRED"

Tue Nov 17 08:55:24 1998

Tue Nov 17 08:55:24 1998

US-08-887-977-9.196

[illegible]

Tue Nov 17 08:55:24 1998

Tue Nov 17 08:55:24 1998

US-08-887-977-9, 1998

LOCUS					PRI	16-JUN-1998
RS125H2	175513 bp	DNA				
DEFINITION	Human DNA sequence from clone 125H2 on chromosome 22q11-12 Contains part of myosin heavy chain gene, EST, CA repeat, STS, GSS, complete sequence.					
ACCESSION	U00497					
NID	F316890					
SOURCE	HTc; myosin heavy chain; repeat polymorphism.					
ORGANISM	human					
REFERENCE	Eukaryotes Metazoa: Chordata; Vertebrata; Mammalia; Eutheria: Primates; Catarrhini; Hominiidae; Homo.					
AUTHORS	1. (bases 1 to 175513)					
JOURNAL	In J. Lloyd D. Submission					
COMMENT	<p>Submitted (13-MAY-1998) E-mail enquiries: humquery@anger.ac.uk Clone requests: clonesrequest@anger.ac.uk During sequence assembly data is compared from overlapping clones. The sequences shown here are those which have been annotated as variations together with notes of their location relative to other features. This variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit segments with only a small overlap as described above. This sequence is the only one available at present, so it has been included in the database according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat regions. In some cases, where the sequence is ambiguous, there is an annotation using the symbol '*'. This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr22/ 125H2 is from the human BAC library described in U.-J. Kim et al. (1996). Genomics 34, 213-218. VECTOR: pBAC10BL.</p>					
FEATURES	Location/Qualifiers					
source	Organism='Homo sapiens' /db_xref='taxon:9606'					
	/chromosome='22'					
	/map='11-12'					
	/map_pos='11-12'					
	complement(211..365)					
repeat_region	/note='MR repeat: matches 192..52 of consensus'					
gene	963..12609[+][R]					
CDS	Join(+563..170,9002..9161.11367..11550.12697..12919.16737..16846.13587..33652.39553..39681.41456..41599.43286..43376.56023..58142.56521..56679.60423..60547.67613..76317.66524..86676.111403..811516.113978..111099.115633)..115716.117513..117689.118673..118708.					

[illegible]

```

repeat_region complement(20212, .20549)
repeat_region /note="LIP1A3 repeat: matches 303, .570 of consensus"
repeat_region complement(20250, .20846)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(21003, .21003)
repeat_region /note="LIP1A3 repeat: matches 573, .362 of consensus"
repeat_region complement(21104, .21286)
repeat_region /note="LIP1A5 repeat: matches 904, .724 of consensus"
repeat_region complement(21193, .21582)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(21243, .21582)
repeat_region /note="LIP1A3 repeat: matches 414, .128 of consensus"
repeat_region complement(21773, .22063)
repeat_region /note="LIP1A3 repeat: matches 903, .570 of consensus"
repeat_region complement(21770, .22282)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(22127, .22509)
repeat_region /note="LIP1A5 repeat: matches 904, .724 of consensus"
repeat_region complement(22516, .22805)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(22829, .23118)
repeat_region /note="LIP1A3 repeat: matches 414, .128 of consensus"
repeat_region complement(23501, .23544)
repeat_region /note="LIP1A3 repeat: matches 300, .7 of consensus"
repeat_region complement(23501, .23544)
repeat_region /note="MIR repeat: matches 56, .98 of consensus"
repeat_region complement(24134, .24436)
repeat_region /note="MIR repeat: matches 5, .201 of consensus"
repeat_region complement(24144, .24436)
repeat_region /note="LIP1A3 repeat: matches 300, .7 of consensus"
repeat_region complement(24724, .24767)
repeat_region /note="LIP1 repeat: matches 56, .98 of consensus"
repeat_region complement(25232, .25515)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(25232, .25515)
repeat_region /note="LIP1 repeat: matches 5021, .4889 of consensus"
repeat_region complement(25356, .25546)
repeat_region /note="MIR repeat: matches 6, .201 of consensus"
repeat_region complement(25356, .25546)
repeat_region /note="LIP1 repeat: matches 300, .1 of consensus"
repeat_region complement(26444, .26576)
repeat_region /note="LIP1 repeat: matches 5021, .4889 of consensus"
repeat_region complement(27192, .27534)
repeat_region /note="LIP1 repeat: matches 345, .1 of consensus"
repeat_region complement(27378, .27772)
repeat_region /note="LIP1 repeat: matches 345, .1 of consensus"
repeat_region complement(27763, .28023)
repeat_region /note="LIP1 repeat: matches 345, .1 of consensus"
repeat_region complement(28023, .28283)
repeat_region /note="LIP1 repeat: matches 345, .1 of consensus"
repeat_region complement(28283, .28543)
repeat_region /note="MIR repeat: matches 345, .1 of consensus"
repeat_region complement(28986, .29246)
repeat_region /note="MIR repeat: matches 345, .1 of consensus"
...
Note: remainder of annotations omitted.

1.69; Score 18; DB 21; Length 173513;
Query Match

```

```

Run Row 17 00:55:24 1998

US-08-887-977-9.ysg

/4b.xref=taxon:9606*
/chromosome=22*
/clocs=2119s
BASE COUNT 44710 a 39974 c 40165 g 49034 t 7280 others
GALVIN

Query Match 1.6% Score 18; DB 13; Length 180163;
Best Local Similarity 100.0%; Pred. No. 1.07e+01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 51664 TGCTGCTGCTGCTGCTGCTG 51681
QY 749 TGCTGCTGCTGCTGCTGCTG 765

RESULT 49
LOCUS AC004479 190000 bp DNA HTG 03-JUL-1998
DEFINITION *** SEQUENCING IN PROGRESS *** Homo sapiens chromosome 4, BAC clone
CO22A4; HTGS Phase 1, 48 unordered pieces.
ACCESSION G1287659
NID
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Vertebrata; Mammalia; Eutheria.
Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 190000)
Stone, N.E., Schmutz, J.J., Cox, D.R. and Myers, R.M.
Direct Submission
2 (bases 1 to 190000)
Stone, N.E., Schmutz, J.J., Cox, D.R. and Myers, R.M.
Direct Submission
Submitted (27-MAR-1998) Department of Genetics, Stanford Human
Genome Center, 95 California Avenue, Palo Alto, CA 94304, USA
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
This sequence is unfinished. It consists of 48 contigs for
the HTGS Phase 1. Only one contig has a complete record in
the database. In some cases the exact length of the
between the contigs are also unknown; these gaps are presented
as runs of N as a convenience only. When sequencing is complete,
the sequence data presented in this record will be replaced.
By a single file containing 1048 bp in length
1049 1195: gap of unknown length
1196 2336: contig of 1121 bp in length
2117 2463: gap of unknown length
3347 3583: gap of unknown length
3584 3693: gap of unknown length
AUTHORS
TITLE
JOURNAL
COMMENT

```


C 553	14	1.3	6828	13	7175302	Genomic RNA of human Rhodospirillum rubrum
C 554	14	1.3	6953	19	QT51016	Partial BCR/CA cancer
C 555	14	1.3	7240	35	799668	Epilinie rhinovirus 1 (
C 556	14	1.3	7277	35	795178	Epilinie rhinovirus 2 (
C 557	14	1.3	7301	35	795178	Epilinie rhinovirus 3 (
C 558	14	1.3	7393	35	774883	Porcine retrovirus GP
C 559	14	1.3	7440	4	R20402	Sequencing of a full-l
C 560	14	1.3	7499	1	N91312	PCP-III plasmid.
C 561	14	1.3	7501	1	N91312	PCP-III plasmid.
C 562	14	1.3	7562	35	785343	PCP-III variable maj
C 563	14	1.3	7790	15	Q84650	Human neuronal calciu
C 564	14	1.3	7808	15	Q84659	Human neuronal calciu
C 565	14	1.3	7972	35	774884	Miniature avian retro
C 566	14	1.3	8020	35	774884	Miniature avian retro
C 567	14	1.3	8060	35	774811	Porcine retrovirus GP
C 568	14	1.3	8196	38	799659	Porcine retrovirus GP
C 569	14	1.3	8209	39	V09700	Porcine retrovirus DN
C 570	14	1.3	8230	39	V09700	Porcine retrovirus DN
C 571	14	1.3	8371	30	766809	Human neurotrophidin
C 572	14	1.3	8601	4	Q24829	APP-1 (C 75108 T)
C 573	14	1.3	8601	4	Q24828	APP-1
C 574	14	1.3	8601	4	Q24828	APP-1
C 575	15	1.3	8760	12	Q13473	BgIII/HpaII fragments
C 576	14	1.3	8797	37	750956	DNA cassette for amph
C 577	14	1.3	9108	16	703104	Plasmid pMUS tk1-1
C 578	14	1.3	9391	1	Q64275	WABURY 1/22 gene res
C 579	14	1.3	9391	1	Q64275	WABURY 1/22 gene res
C 580	14	1.3	9501	1	Q64275	WABURY 1/22 gene res
C 581	14	1.3	9739	37	758490	Sequence corresp to
C 582	14	1.3	9739	37	758490	Sequence corresp to
C 583	14	1.3	9739	37	758490	Sequence corresp to
C 584	14	1.3	9746	37	758450	Human L1Mmudeficlen
C 585	14	1.3	10079	35	792317	Xiv-1(Mmu)ov protein
C 586	14	1.3	10079	35	792317	Xiv-1(Mmu)ov protein
C 587	14	1.3	10937	37	750855	Human L1Mmudeficlen
C 588	14	1.3	10937	37	750855	Human L1Mmudeficlen
C 589	14	1.3	11283	35	784841	Human breast and ovar
C 590	14	1.3	11388	36	769701	Human breast cancer
C 591	14	1.3	11478	33	779805	Full length potato at
C 592	14	1.3	11478	33	779805	Full length potato at
C 593	14	1.3	11716	36	791087	Bovine neurotrophidin
C 594	14	1.3	12151	11	Q65137	Rice starch branching
C 595	14	1.3	15144	29	747115	Mouse bone morphogen
C 596	14	1.3	15144	29	747115	Mouse bone morphogen
C 597	15	1.3	18216	24	774408	RCMV Toledo strain no
C 598	15	1.3	20710	12	710681	Kaposi's sarcoma asso
C 599	14	1.3	37895	24	734137	Gene for biosynthetic
C 600	14	1.3	37913	10	Q51238	Zoonotic enteric acid
C 601	14	1.3	37913	10	Q51238	Zoonotic enteric acid
C 602	14	1.3	4937	39	V05387	The saccharon biosynth
C 603	14	1.3	53526	40	794101	Human DPB1 gene.

[illegible]

C 555	1.1	252.16	Q86075	Polypeptide chain	1.13e+03
C 556	1.1	256.32	T31009	Erythropoietin 3 / hyp	1.13e+03
C 557	1.1	262.13	Q75167	ALL-1/AF-9 chimeric c	1.31e+03
C 558	1.1	284.42	T26729	Human gene signature	1.31e+03
C 559	1.1	284.42	T26729	Human gene signature	1.31e+03
C 560	1.1	307.38	T00043	Hepatitis C virus HA	1.31e+03
C 561	1.1	317.38	Q50571	DNA associated with h	1.31e+03
C 562	1.1	324.35	T16564	Cancer detection prob	1.31e+03
C 563	1.1	324.35	T16564	Cancer detection prob	1.31e+03
C 564	1.1	340.32	Q14085	Human gene signature	1.31e+03
C 565	1.1	363.32	T26639	Human gene signature	1.31e+03
C 566	1.1	363.3	Q61303	Human brain expressed	1.31e+03
C 567	1.1	370.3	Q69237	Human brain expressed	1.31e+03
C 568	1.1	370.3	Q69237	Human brain expressed	1.31e+03
C 569	1.1	380.8	Q60264	Human brain expressed	1.31e+03
C 570	1.1	395.8	Q48713	ADP-Glucose-Pyrophos	1.31e+03
C 571	1.1	396.40	Q90999	3 portion of clone B	1.31e+03
C 572	1.1	396.40	Q90999	3 portion of clone B	1.31e+03
C 573	1.1	399.8	Q61271	Human brain expressed	1.31e+03
C 574	1.1	404.8	Q67078	Human brain expressed	1.31e+03
C 575	1.1	406.18	T19293	Human gene signature	1.31e+03
C 576	1.1	417.38	Q00456	Mouse preprocalcitat	1.31e+03
C 577	1.1	427.38	Q00456	Mouse preprocalcitat	1.31e+03
C 578	1.1	465.39	T55517	Hamster oral papillom	1.31e+03
C 579	1.1	475.37	T79850	Flea saliva protein	1.31e+03
C 580	1.1	476.36	T94932	Human endometrial spe	1.31e+03
C 581	1.1	480.38	T94932	Human endometrial spe	1.31e+03
C 582	1.1	540.39	T39802	Human clonin 3/3 gene	1.31e+03
C 583	1.1	576.31	T15601	Hepatitis C virus like	1.31e+03
C 584	1.1	591.37	Q70794	PAHox coat protein 26	1.31e+03
C 585	1.1	592.8	Q45053	Barley subtilisin inh	1.31e+03
C 586	1.1	617.38	Q13267	Human gene signature	1.31e+03
C 587	1.1	627.39	T53267	S. lepdodophylla trhea	1.31e+03
C 588	1.1	632.33	T60535	HIV-1 5' long termina	1.31e+03
C 589	1.1	643.40	Q10393	Sheath virus nucleic	1.31e+03
C 590	1.1	643.40	Q10393	Sheath virus nucleic	1.31e+03
C 591	1.1	673.38	Q50573	DNA associated with h	1.31e+03
C 592	1.1	685.40	Q101590	Sheath virus nucleic	1.31e+03
C 593	1.1	688.40	Q101800	Sheath virus nucleic	1.31e+03
C 594	1.1	712.3	Q104309	At. pyrocl. oclary mem	1.31e+03
C 595	1.1	712.3	Q104309	At. pyrocl. oclary mem	1.31e+03
C 596	1.1	720.39	T143447	Green fluorescent pro	1.31e+03
C 597	1.1	720.39	T14336	Wild type green fluor	1.31e+03
C 598	1.1	720.39	T14343	Green fluorescent pro	1.31e+03
C 599	1.1	720.39	T14343	Green fluorescent pro	1.31e+03
C 600	1.1	720.39	T14345	Green fluorescent pro	1.31e+03
C 601	1.1	720.39	T14344	Green fluorescent pro	1.31e+03
C 602	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 603	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 604	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 605	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 606	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 607	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 608	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 609	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 610	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 611	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 612	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 613	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 614	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 615	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 616	1.1	720.39	T14346	Green fluorescent pro	1.31e+03
C 617					

[illegible]

US-08-887-977-9.rmg

Tue Nov 17 08:55:26 1998

Biotin synthetase enzy 1.31e+03
 Rat 5-HT_{1B} receptor 1.31e+03
 Correct A. thaliana p 1.31e+03
 Incorect A. thaliana p 1.31e+03
 Rad A gene encoding g 1.31e+03
 Human cytochrome P450 1.31e+03
 Brevibacterium lactof 1.31e+03
 Papillomavirus E2 bin 1.31e+03
 Collagen binding prot 1.31e+03
 Arabidopsis floral me 1.31e+03
 Human brain specific 1.31e+03
 Plasmid pCMVgfp1 exp 1.31e+03
 C. glutamicum pC19.1 1.31e+03
 Stearoyl-ACP-desatur 1.31e+03
 Survival motor neuron 1.31e+03
 cDNA encoding transfo 1.31e+03
 Human J59-41 secreted 1.31e+03
 Cotton fibre specific 1.31e+03
 Cotton fibre cell-ape 1.31e+03
 Wt1a rativa cytochr 1.31e+03
 Analogue of cDNA codi 1.31e+03
 Novel plasmogen acti 1.31e+03
 Analogue of cDNA codi 1.31e+03
 Penicillium requorist 1.31e+03
 Mch4 coding sequence 1.31e+03
 Human protease PMH-1 1.31e+03
 Human L-asparaginase 1.31e+03
 Human L-asparaginase 1.31e+03
 Clone pACYC encoding 1.31e+03
 P. gingivalis cell su 1.31e+03
 Human RNA-binding pro 1.31e+03
 Brevibacterium flavum 1.31e+03
 Cyclochrome P450 X9 c 1.31e+03
 Oazyl-CoA decarboxyl 1.31e+03
 Buprl protease inhibi 1.31e+03
 Full length cDNA clon 1.31e+03
 pPR4 gene encoding p1 1.31e+03
 Fragment of tobacco p 1.31e+03
 Human segment cDNA enc 1.31e+03
 Tumour suppressor gen 1.31e+03

US-08-887-977-9.rmg

Tue Nov 17 08:55:26 1998

Rat vesicle associate 1.31e+03
 cDNA for wild type hu 1.31e+03
 Human K₁ channel cDN 1.31e+03
 Rabphilin-3A cDNA 1.31e+03
 Clone 4-33 encoding A 1.31e+03
 Rhodospirillum rubrum 1.31e+03
 Rhodospirillum rubrum 1.31e+03
 Human calpastatin cDN 1.31e+03
 Mouse 2-5A-dependent 1.31e+03
 Promoter from B. flavu 1.31e+03
 R55 Cys² antitense tr 1.31e+03
 Human metallothionein 1.31e+03
 IS1096 insertion sequ 1.31e+03
 Rat kynurenine aminot 1.31e+03
 Tumour rejection anti 1.31e+03
 Homoserine dehydrogen 1.31e+03
 cDNA of clone GEM178 1.31e+03
 Anthr specific promo 1.31e+03
 Potato viral gene with 1.31e+03
 Retinoid X receptor 1 1.31e+03
 Protein kinase (CKIIa) 1.31e+03
 Murine TRAC K3 isofor 1.31e+03
 Human cytochrome P450 1.31e+03
 Streptococcus pneumonia 1.31e+03
 Enodermis M- and N-prot 1.31e+03
 DNA encoding tumour a 1.31e+03
 Rat choline kinase ge 1.31e+03
 K. pflori cytoplasmic 1.31e+03
 Recombinant bovine ed 1.31e+03
 Human metallothionein 1.31e+03
 Cyclic-GMP stimulated 1.31e+03
 Plectin ent-karum B 1.31e+03
 Sequence of gene for 1.31e+03
 Human CRIPRO-related 1.31e+03
 Class A starch branch 1.31e+03
 Drosophila p7056K gen 1.31e+03
 Reast acyl-coenzyme A 1.31e+03
 Human cytochrome P450 1.31e+03
 S. capsulata prolyi o 1.31e+03
 Human cytochrome P450 1.31e+03
 Gene encoding AmpiOL 1.31e+03
 Bovine protein diuap 1.31e+03

US-08-887-977-9.rmg

Tue Nov 17 08:55:26 1998

PT-WANBH virus BHC-11 1.31e+03
 Human NF-AT120, C sub 1.31e+03
 Unique 2.9 kb genomic 1.31e+03
 Tactitoxyl amino acid 1.31e+03
 Clone encoding rat ov 1.31e+03
 Cry2B gene which enc 1.31e+03
 Clone 25-1 encodes hu 1.31e+03
 H2L17/77 cDNA encodi 1.31e+03
 Human G-protein coupl 1.31e+03
 Cancer suppressor gen 1.31e+03
 Class A starch branch 1.31e+03
 Saccharomyces cerevis 1.31e+03
 Saccharomyces cerevis 1.31e+03
 Endodermis parietal Bu 1.31e+03
 Human disc large 1 g 1.31e+03
 Human prostatic protei 1.31e+03
 3a3' sequence of the 1.31e+03
 HEK coding sequence 1.31e+03
 Bovine lysosomal alph 1.31e+03
 Bovine lysosomal alph 1.31e+03
 Delta(4)-5-oxosterol 1.31e+03
 Arabidopsis thaliana g 1.31e+03
 Arabidopsis thaliana g 1.31e+03
 Human G-protein coupl 1.31e+03
 Human metabotropic gl 1.31e+03
 KARS coding sequence 1.31e+03
 Pyrococcus furiosus D 1.31e+03
 Rabbit poly-immunoglob 1.31e+03
 Human embryonal Kinas 1.31e+03
 HER-2/neu oncogene 1.31e+03
 Mouse beta 2 integrin 1.31e+03
 Human alpha-D subunit 1.31e+03
 Pot1a1 nucleotide se 1.31e+03
 Sequence encoding Inc 1.31e+03
 Tomato cr2 gene 1.31e+03
 Endogen Protein tyros 1.31e+03
 Varicella zoster viru 1.31e+03
 Sequence of LAV/HTLV 1.31e+03
 DNA clone of lambdaDUC 1.31e+03
 Rat inducible phosph 1.31e+03
 Human Stem Cell Facto 1.31e+03
 Human glycogen phosph 1.31e+03

US-08-887-977-9.rmg

Tue Nov 17 08:55:26 1998

Human inducible nitri 1.31e+03
 Singapore horseshoe c 1.31e+03
 Stem cell leukemia (1.31e+03
 Sequence of precursor 1.31e+03
 Cytoskeleton encoding hig 1.31e+03
 Sequence encoding hig 1.31e+03
 Apoptosis inducing pr 1.31e+03
 Drosophila melanogast 1.31e+03
 Cytoskeleton encoding 1.31e+03
 Cytoskeleton encoding 1.31e+03
 Human retinoblastoma 1.31e+03
 Human multigene resis 1.31e+03
 Sequence of human mul 1.31e+03
 R6 promoted cytomegal 1.31e+03
 Human papillomavirus 1.31e+03
 Neuronal nitrogen mon 1.31e+03
 Rag A antigen of HeLa 1.31e+03
 Sequence of boyl cDN 1.31e+03
 Human additional sex 1.31e+03
 Human additional sex 1.31e+03
 Impatiens Necrotic Sp 1.31e+03
 Sequence of boyl cDN 1.31e+03
 5.3 Kb EcoRI fragmen 1.31e+03
 Neuronal apoptosis in 1.31e+03
 Sequence encoding enz 1.31e+03
 Shodococcus chondroch 1.31e+03
 Human SPT6 cDNA 1.31e+03
 Neuronal apoptosis in 1.31e+03
 Plasmid pPRMD7 expre 1.31e+03
 Picea mariana thallus 1.31e+03
 Mouse Friend virus su 1.31e+03
 Hybrid vector pSP-MDR 1.31e+03
 Human neuronal calciu 1.31e+03
 Arabidopsis thaliana 1.31e+03
 DNA associated with h 1.31e+03
 H6 promoted cytomegal 1.31e+03
 Lactobacillus bulgari 1.31e+03
 Genomic DNA LF-ACC2 1.31e+03
 ACCTA cDNA 1.31e+03
 Wheat acetyl-CoA carb 1.31e+03
 Glucocorticoidase ge 1.31e+03
 Cell cycle checkpoint 1.31e+03
 Human cyclin D1 1.31e+03
 FelV F6A provirus clo 1.31e+03

[illegible]

Tue Nov 17 08:55:26 1998

UG-08-887-977-9.mpg

[illegible]

Page 34

US-08-087-977-9-1708

[illegible]

US-08-887-977-9-ENG

PAGE 32

US-08-887-977-9.mxd

CDNA.
al transduction;

— ПОРЯДОК ВНЕ

- useful to identify:

8, etc.

Human G-protein chemokine receptor 1 (GPR15) is a G-protein coupled receptor (GPCR) involved in chemotaxis and immune response. It is a transmembrane protein with seven transmembrane domains. The extracellular domain contains a binding site for chemokines, and the intracellular domain contains a binding site for G-proteins. The G-protein involved in the signaling pathway is G_i.

incorporated into a vector

to underexpression

C. 313 G: 407 T:

5; Length 1414;

```
0; Indels 0;
```

RESULT 11
 ID T90117 standard: cDNA: 1477 BP.
 AC T90117 (first entry)
 DE cDNA for human CCR5.
 KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus type 1; HIV-1; HIV-2;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder; ss.
 KW Homo sapiens. Location/Qualifiers
 FT Key 240..1398
 FT CDS /*tag= a
 FT W09712019-A2.
 FT 14-SEP-1997. B00023
 FT 24-SEP-1997. EP-871012.
 FT 06-AUG-1996. EP-871012.
 FT 01-MAR-1996. EP-871021.
 PA (EURO-) EUROSCREEN SA.
 PI Libert F, Farnetier M, Samson M, Vassart G;
 DR P-PSDB: 479839/44.
 DR P-PSDB: T90117.
 FT Active and inactive forms of human CC chemokine receptor CCR-5 -
 FT useful to diagnose, prevent and/or treat inflammatory disorders,
 FT autoimmune disease and viral infection
 FT Claim 18: FIG 1b-c: 94pp: English
 CC The present sequence encodes human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC (MIP-1, MIP-1 beta, but not by monocyte chemoattractant protein 1 (MCP-1)).
 CC MIP-2, MIP-3, MIP-4, MIP-5, MIP-6, MIP-7, MIP-8, MIP-9, MIP-10, MIP-11,
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis, fibrosis and
 CC glomerulonephritis.
 CC CCR5 is a receptor for HIV-1 and HIV-2.
 CC Psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 CC Sequence 1477 BP; 374 A; 320 G; 431 T;
 SQ Query Match 1.5%; Score 17; DB 37; Length 1477;
 Best Local Similarity 100.0%; Pred. No. 3.85e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1144 agaagttcgaactac 1160
 QY 935 AGAGTTTCGAAGTAC 951

QY 935 AGAGTTTCGAAGTAC 951
 RESULT 10
 ID T90118 standard: cDNA: 1442 BP.
 AC T90118 (first entry)
 DE cDNA for inactive human CCR5.
 KW Inactive human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW predisposition; resistance; diagnosis; treatment; prevention;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder; ss.
 KW Homo sapiens. Location/Qualifiers
 FT Key 240..1398
 FT CDS /*tag= a
 FT W09712019-A2.
 FT 04-SEP-1997. B00023
 FT 24-SEP-1997. EP-871012.
 FT 06-AUG-1996. EP-871012.
 FT 01-MAR-1996. EP-871021.
 PA (EURO-) EUROSCREEN SA.
 PI Libert F, Farnetier M, Samson M, Vassart G;
 DR P-PSDB: 479839/44.
 DR P-PSDB: T90117.
 FT Active and inactive forms of human CC chemokine receptor CCR-5 -
 FT useful to diagnose, prevent and/or treat inflammatory disorders,
 FT autoimmune disease and viral infection
 FT Claim 18: FIG 1b-c: 94pp: English
 CC The present sequence encodes human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC (MIP-1, MIP-1 beta, but not by monocyte chemoattractant protein 1 (MCP-1)).
 CC MIP-2, MIP-3, MIP-4, MIP-5, MIP-6, MIP-7, MIP-8, MIP-9, MIP-10, MIP-11,
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis, fibrosis and
 CC glomerulonephritis.
 CC CCR5 is a receptor for HIV-1 and HIV-2.
 CC Psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 CC Subjects that express the inactive receptor have a predisposition,
 CC to develop inflammatory diseases.
 CC Sequence 1442 BP; 343 A; 314 G; 422 T;
 SQ Query Match 1.5%; Score 17; DB 37; Length 1442;
 Best Local Similarity 100.0%; Pred. No. 3.85e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1112 agaagttcgaactac 1128
 QY 935 AGAGTTTCGAAGTAC 951

RESULT 12
 ID T90107 standard: cDNA: 1526 BP.
 AC T90107 (first entry)
 DE Human NF-1 gene clone KF361-2.
 KW Alzheimer's disease; NF-1; human; mouse; KF361 gene; antibody; ss.
 KW Homo sapiens.
 KW 109213495-A.
 KW 10-NOV-1996. JP-345659.
 KW 07-DEC-1995. JP-345659.
 PA (SDSU) SUMITOMO SEIYAKU KK.
 PI WPI: 97-46507/43.
 DR P-PSDB: 479839/44.
 DR P-PSDB: T90117.
 FT Active and inactive forms of human CC chemokine receptor CCR-5 -
 FT useful to diagnose, prevent and/or treat inflammatory disorders,
 FT autoimmune disease and viral infection
 FT Claim 18: FIG 1b-c: 94pp: English
 CC The present sequence encodes human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC (MIP-1, MIP-1 beta, but not by monocyte chemoattractant protein 1 (MCP-1)).
 CC MIP-2, MIP-3, MIP-4, MIP-5, MIP-6, MIP-7, MIP-8, MIP-9, MIP-10, MIP-11,
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis, fibrosis and
 CC glomerulonephritis.
 CC CCR5 is a receptor for HIV-1 and HIV-2.
 CC Psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 CC Sequence 1526 BP; 436 A; 269 C; 315 G; 506 T;
 SQ Query Match 1.5%; Score 17; DB 34; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 3.85e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1015 tggcctctctatataaa 1031
 Cp 205 TGGCCTCTCTATATAAA 189
 RESULT 13
 ID T99542 standard: cDNA: 1557 BP.
 AC T99542 (first entry)
 DE Human chemokine receptor MMR-CCR cDNA.
 KW human; monocyte; macrophage; chemotaxis; human; infection;
 KW inflammation; proliferative disease; cardiovascular disease;
 KW tumour; rheumatoid arthritis; alveolitis; atherosclerosis;
 KW chronic granulomatous disease; asthma; myasthenia gravis;
 KW septic shock; Chediak-Higashi syndrome; therapy; diagnosis; ss.
 KW Homo sapiens. Location/Qualifiers
 FT Key 36..1034
 FT CDS /*tag= a
 FT W09712125-A2.
 FT 06-NOV-1997. U06993.
 FT 25-APR-1997. U06993.

RESULT 13
 ID T99542 standard: cDNA: 1557 BP.
 AC T99542 (first entry)
 DE Human chemokine receptor MMR-CCR cDNA.
 KW human; monocyte; macrophage; chemotaxis; human; infection;
 KW inflammation; proliferative disease; cardiovascular disease;
 KW tumour; rheumatoid arthritis; alveolitis; atherosclerosis;
 KW chronic granulomatous disease; asthma; myasthenia gravis;
 KW septic shock; Chediak-Higashi syndrome; therapy; diagnosis; ss.
 KW Homo sapiens. Location/Qualifiers
 FT Key 36..1034
 FT CDS /*tag= a
 FT W09712125-A2.
 FT 06-NOV-1997. U06993.
 FT 25-APR-1997. U06993.

```

Matches      17; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db 1990 tgcctcattataaaaa 2006
Cp   205 TGGCTCTTTATATAAAA 189

RESULT 16
AC Q02068 standard; cDNA; 2080 bp.
AD T80098;
DE DT 25-NOV-1997 (first entry)
DE KF-1 gene;
DE MW Alzheimer's disease; KP-1; human; mouse; KP361 gene; antibody; aa.
FH KJ040188;
FH KW Human apofAin. Location/Qualifiers
FT FT 1..1641 /tag= a
FT TT J09215495-A.
PD PD 19-AUG-1997.
PD PP 20-JUN-1986; I81514.
PR PR 07-DEC-1995; JP-345659.
PR PS P-PSB: W26164.
RA RA WP0197-45037/v3. SEITAKU KK.
RE RE P-PSB: W26164.
DE DE P-PSB: W26164.
PT PT A brain-specific expression gene - used in the diagnosis of
PT Alzheimer's disease. Japan.
PT This sequence represents the human
PT This sequence represents the human Kf361 gene. This sequence contains the
PT open reading frame represented by T80102. This sequence, T80099, and
PT T80100 all represent cDNA sequences of the invention. These sequences are
PT identical except for one base pair difference. The reading frame product of
PT T80103 and T80101 is identical. Open reading frames product of
PT T80103 and T80102 are identical. These sequences are useful for the diagnosis of
PT antibodies specific for the proteins are useful for the diagnosis of
PT Alzheimer's disease.
SQ SQ Sequence 2080 BP; 366 C; 424 G; 664 T;

Best Match
Query Local Similarity 100.0%; Pred. No. 3.85e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1573 tgcctcattataaaaa 1589
Cp   205 TGGCTCTTTATATAAAA 189

RESULT 17
ID Q02068 standard; DNA; 2135 bp.
AC Q02069;
AC Q02069; 1989 (first entry)
DE DT 25-NOV-1997 (first entry)
DE KF-1 gene;
DE MW Human muscarinic acetylcholine m2 receptor gene
FH KJ040188;
FH KW Muscarinic acetylcholine recepior; drug screening; probes; m2; aa.
```

```
PH Location/Qualifiers
FT   120..1517
FT   cds
FT   /tag = a
FT   /product=MAR subtype m2
FT   misc_signal
FT   /tag = b
FT   /label-splice acceptor site
FT   /note=defines the 5' end of the exon
FT   poly_a_site
FT   /note=2096 C
FN   OS2741971-A.
FN   14-MAR-1989.
FF   08-SEP-1989; 741971.
FF   08-SEP-1989; 741971.
FA   (USRS) Nat Inst of Health.
FW   WPI; 59-16545/22.
DR   P-RSDS; P96203.
DR   P-RSDS; P96203.
PT   Description: nuclear-stylectobaline receptors -
PT   for drug screening and diagnostic use.
PT   Disclosure: P; English.
CC   The sequence may be useful for synthesis of hybridization probes for
CC   c-myc gene. The genes are cloned by screening 725 cDNA libraries
CC   from rat brain with a c-myc probe. A c-myc probe was also used to
CC   strand of porcine brain (ml) cDNA modified at positions 5', 3g and 51).
CC   Identifying cDNA clones by hybridization with BamHI or XhoI digests of
CC   plasmid DNA from 12 cultures contg. up to 50000 independent clones.
CC   The c-myc gene was identified as a single copy.
CC   single hybridisation band are identified, and isolating individual clones
CC   by colony hybridization.
CC   Stable cell line produced by transfecting Chinese hamster ovary cells
CC   with the c-myc gene construct. The gene inserts
CC   See also G82048-N82067 and N8008.
SQ   Sequence 2135 BP; 629 A; 462 C; 448 G; 596 T;
Query Match          15%      Score 17; DB: 2; Length 2135;
Matchase 17; Conservative      And N.
                                O; Mismatches 0; Indels 0; Gaps 0;
Db    393 ggaactgggtggagaa 409
Qt    969 GGACCTGTCGGTGTGA 985
RESURF 19
ID AC T85161 standard; cDNA; 3383 BP.
ID AC T85161;
DE DE 14-DER-1987 (first entry);
DE Human chemokine receptor 89c cDNA.
DE Tumor necrosis factor alpha receptor; chemokine; rhinovirus; rheumatoid arthritis;
DE tumour; asthma; viral infection; AIDS; inflammation; leukocyte trafficking;
DE autoimmune disease; therapy; diapedesis; leukocyte trafficking;
```

07-DDC-1995; JP-345659.
 PR (SDU) SUMITOMO SEIZAKU KK.
 WI: 97-465073/43.
 CC A brain-specific expression gene - used in the diagnosis of
 PT Alzheimer's disease
 PP Alzheimer's disease
 PS Claim 6, Page 12, 13pp; Japanese.
 CC This sequence represents the open reading frame of the KP-1 gene.
 CC This sequence represents the open reading frame of the KP-2 gene.
 CC invention, and are all KP-1 genes. The genes contain the open reading
 CC frames represented by this sequence, T80101 and T80103. These genes,
 CC their products and antibodies specific for the proteins are useful for
 CC the diagnosis of Alzheimer's disease.
 CC Sequence 1641 bp; 489 A; 300 C; 351 G; 495 T;
 SEQ Sequence 1641 bp; 489 A; 300 C; 351 G; 495 T;
 Query Match 1.5%; Score 17; DB 34; Length 1641;
 Best Local Similarity 100.0%; Pred. No. 3.85e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1573 tgcctctctatcaasaa 1589
 Cp 205 TGCCCTCTTATTAAMA 189
 RESULT 15
 ID T80103 standard; cDNA: 205B BP.
 DT 25-NOV-1997 (first entry)
 DE Human KP-1 gene open reading frame.
 DE Alzheimer's disease: KP-1; human; mouse; KP361 gene; antibody: m5.
 DE Accession: J09215494cna.
 DE J09215494cna.
 PD 13-AUG-1997.
 PD 20-JUN-1996; 181514.
 PR 07-DDC-1995; JP-345659
 PR (SDU) SUMITOMO SEIZAKU KK.
 WI: 97-465073/43.
 CC A brain-specific expression gene - used in the diagnosis of
 PT Alzheimer's disease
 PP Alzheimer's disease
 PS Claim 6, Page 12, 13pp; Japanese.
 CC This sequence represents the open reading frame of the human KP-1 gene.
 CC This sequence represents the open reading frame of the human KP-2 gene.
 CC invention, and are all KP-1 genes. The genes contain the open reading
 CC frames represented by this sequence, T80098-T80100 all represent cDNAs of
 CC the genes, their products and antibodies specific for the proteins are useful for
 CC the diagnosis of Alzheimer's disease.
 CC Sequence 205B bp; 584 A; 388 C; 477 G; 609 T;
 SEQ Sequence 205B bp; 584 A; 388 C; 477 G; 609 T;
 Query Match 1.5%; Score 17; DB 34; Length 205B;
 Best Local Similarity 100.0%; Pred. No. 3.85e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

US-08-087-977-9 .rnx

Nov 17 08:55:26 1998

OS Homo sapiens.
FF Key Location/Qualifiers
FF FR cd5 120..1517
FF FT /*tag= a
FF FT /*product=VAR subtype a2
FF FT misc_signal 74
FF FT /*tag= b
FF FT /*label=splice acceptor site
FF FT /*note=defines the 5' end of the exon
FF FT poly_a_site 71038
FF FT /*tag= c
FF FT US741971-A.
FF FR 14-WK-1989. 241971
FF FR 18 SEP-1989. 241971
FF FR 08 SEP-1989. 241971
FF PA (GSS) Nat Inst of Health.
FF FR WPI: 89-165452/22.
FF DR P-PDB9; R02201.
FF DR P-PDB9; R02201.
FF FT muscarinic acetylcholine receptors -
FF FT for drug screening and diagnostic use.
FF PS Disclosure: P. English.
CCC The sequence may be useful for synthesis of hybridization probes for
CCC diagnostic use. The genes are cloned by screening a rat cerebral cortex
CCC cDNA library (from Dr. J. W. Fawcett) with BamHI and XbaI digests of
CCC strand of porcine brain m1 cDNA (modified at positions 538 and 531)
CCC identifying cDNA clones by hybridisation with BamHI or XbaI digests of
CCC plasmid DNA from 12 cultures contg. up to 50000 independent clones;
CCC reselecting until a pure contg. less than 5000 indep. clones with a
CCC cDNA library hybridisation
CCC by colony hybridisation are identified, and isolating individual clones
CCC Stable cell lines are produced by transfecting Chinese hamster ovary cells
CCC (CHO-K1) with various pCD vectors contg.
CCC See also Q02068-Q02067 and Q00086.
CCC The gene inserts.
FF SQ Sequence 2135 bp; 235 A; 442 C; 596 T;
Query Match 1.5%: Score 17; DB 1; Length 2135;
Best Local Similarity 100.0%: Pred. No. 3.85e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 393 ggaactgtgtgtgtga 409
Oy 969 ggaactgtgtgtgtgtga 985
RESULT 18
ID N92069 standard: DNA; 2135 bp.
AC AC N92069;
DB 27-SEP-1989 (first entry)
DB 08 SEP-1989 (second entry)
OS Muscarinic acetylcholine receptor; drug screening; probes; m2; as.
FF FR Homo sapiens.

```


[illegible]

119-08-997-977-9

[illegible]

US-08-887-977-9.rmg

CC proadherent properties for circulating T cells and monocytes. A secreted
 CC or soluble form, consisting of the chemokine domain and the stalk region,
 CC membrane bound form of CX3C chemokine may be a potent regulator of
 CC circulating leukocytes, and thus may be involved in various inflammatory
 CC diseases, e.g. arthritis. The soluble form may be used as a regulator of
 CC chemokine production, especially in conditions of compromised immune response.
 CC inflammatory disorders as multiple sclerosis, and other pathologies
 CC involving neurogenic inflammation. The products can be used to diagnose
 CC disorders associated with CX3C chemokine misregulation. They can also be
 CC used in the treatment of conditions associated with abnormal
 CC chemokine production, especially in conditions of chronic inflammation.
 CC Abnormal proliferation, regeneration, degeneration, and atrophy may be
 CC modulated by appropriate therapeutic treatment using the products. The
 CC products can also be used in screening assays.

Query Match 1.44; Score 16; DB 36; Length 1654;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1433 attccacagagcca 1448
 |||||

Cp 165 attccacagagcca 150
 |||||

RESULT 33

ID T77791 standard; DNA: 1809 BP.
 AC T77791: 1809 (first entry)
 DT 13-OCT-1989 (first entry)
 KW 6-hydroxynicotinic acid mono-oxygenase.
 KW enzyme: 6-hydroxynicotinic acid mono-oxygenase; 6-HNMO.
 OS Pseudomonas fluorescens strain TMS.
 CC 405-1566
 CC /tag= a
 CC /product= 6-hydroxynicotinic acid mono-oxygenase
 CC /note= "claim 2"
 FT mat_peptide 406..1566
 FT J09121864-A.
 FT 13-MAY-1997.
 FT 08-NOV-1997.
 FT 08-NOV-1997.
 FT (HAIN) COSMO OIL CO LTD.
 PA (COSM-) COSMO SGO KENYUSHO KK.
 DR WPI: 97-314224/29.
 DR SDB: 100018.
 PT DNA encoding 6-hydroxy-nicotinic acid mono-oxygenase - useful for

US-08-887-977-9.rmg

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 940 attctatgcatcaact 955
 |||||
 QY 343 ATTCTATGCACTACT 358

RESULT 35

ID T90132 standard; cDNA: 2499 BP.
 AC T90132: 2499 (first entry)
 DT 10-MAR-1993 (revised)
 DT 13-OCT-1989 (first entry)
 KW Synthetic human 5-lipoxygenase gene.
 KW cDNA: lipoxygenase; biosynthesis of leukotrienes;
 OS Homo sapiens (Human)
 FT CDS 35..1962
 FT misc_feature 1..35
 FT /tag= a
 FT /note= "attached sequence"
 FT EP-325773-A.
 PD 02-AUG-1989.
 PD 28-DEC-1997.
 PA (RISB) Japan Tobacco Inc.
 DR WPI: 89-21819/31.
 DR P-PSDB: P90730.
 PT - used for mass prodn. of 5-lipoxygenase which catalyses
 PT biosynthesis of leukotriene cpds.
 PS Claim 1; fig 2. lipo: English.
 CC Synthetic human 5-lipoxygenase gene (see corresp. P90730). This
 CC cDNA encodes a protein of 562 amino acids. The sequence
 CC is added on at 1900-1903 the sequence is TGA, TAA or TAG.
 CC Sequence 2499 BP: 593 A; 715 C; 662 G; 527 T; 2 Others:

Query Match 1.44; Score 16; DB 1; Length 2499;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1597 caagctctcagcttc 1612
 |||||

QY 996 CAACTCTCAGCTTC 1011

RESULT 36

ID T05136 standard; DNA: 2500 BP.
 AC T05136: 2500 (first entry)
 DT 04-JAN-1991 (first entry)
 DE Sequence encoding human 5-lipoxygenase.
 KW E.coli expression vector; leukotriene A4; lipoxins; ds;

Tue Nov 17 08:55:26 1998

US-08-887-977-9.rmg

PT Industrial production of 2,5-di-hydroxy-pyridine
 CC Claim 2; Page 10-12; 16pp; Japanese.
 CC This sequence encodes 6-hydroxynicotinic mono-oxygenase (6-HNMO) isolated
 CC from Pseudomonas fluorescens strain TMS. The product is a dimeric enzyme
 CC transformed with it, are used to produce 6-HNMO. 6-HNMO is used for the
 CC industrial production of 2,5-di-hydroxypyridine.

Query Match 1.44; Score 16; DB 32; Length 1809;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 117 tcaatctcagcttc 132
 |||||

Cp 67 TCAATCTCAGCTTC 52
 |||||

RESULT 34

ID Q26053 standard; cDNA: 1836 BP.
 AC Q26053:
 DT 14-DEC-1992 (first entry)
 KW Aspergillus oryzae.
 KW saccharification; starch; sugar; brewing; fermentation;
 OS Aspergillus oryzae.
 PN J0418683-A.
 PN 09-OCT-1990; JP-269416.
 PP 09-OCT-1990; JP-269416.
 PA (JOUR-) JOURNAL OF BIOCHEMISTRY.
 DA (JOUR-) JOURNAL OF BIOCHEMISTRY.
 PT New DNA sequence isolated from Aspergillus oryzae - contains
 PT promoter and/or protein coding region of glucanase
 PS Disclosure: Fig 1; 10pp; Japanese. useful for the process of
 CC this sequence codes glucanase. The product is a dimeric enzyme
 CC transformed with it, are used to produce 6-HNMO. 6-HNMO is used for the
 CC industrial production of 2,5-di-hydroxypyridine. When used in brewing
 CC sugar. The A. oryzae/PGAL transformant, when used in brewing
 CC decrease the amount of koji and of sake less, and cost of materials.
 CC Using the 8. cerevisiae transform allows alcohol fermentation
 CC from malted barley and malted wheat. The product is a dimeric enzyme
 CC cDNA library prepared in lambda gt10. Glucanase DNA
 CC isolated by colony hybridization and cloned into a 8. cerevisiae
 CC expression vector. Transformants may be used for direct alcohol
 CC fermentation after fixation on a column wall or beads, substrates
 CC and substrates are applied and ethanol can be eluted.
 CC See also Q26054-6
 CC Sequence 1836 BP: 399 A; 516 C; 493 G; 428 T;

Query Match 1.44; Score 16; DB 4; Length 1836;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;

Tue Nov 17 08:55:26 1998

US-08-887-977-9.rmg

OS Homo sapiens.
 FR Key Location/Qualifiers
 FT cds 38..2056
 FT /tag= a
 PD EP-384750-A.
 PD 29-AUG-1990.
 PP 22-FEB-1989; JP-040206.
 PP 22-FEB-1989; JP-040206.
 PT (RISB) Japan Tobacco Inc.
 PT Matsuyama T. Nishimura M. Nakamura M. Noma M;
 WPI: 90-26217/35.
 DR P-PSDB: R06552.
 PT Prodn. of human 5-lipoxygenase - in large ants. By transformation
 PT of cultured cells, with vector contg. DNA encoding enzyme and
 PT culturing.
 PS Disclosure: Fig 1; 16pp; English.
 CC Sequence may be used to transform an E.coli expression system for
 CC large scale production of human 5-lipoxygenase, useful in formation
 CC of leukotrienes and lipoxins.
 CC Sequence 2500 BP: 708 C; 674 G; 527 T;

Query Match 1.44; Score 16; DB 1; Length 2500;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1597 caagctctcagcttc 1612
 |||||

QY 996 CAACTCTCAGCTTC 1011

RESULT 37

ID M7805 standard; DNA: 2574 BP.
 AC M7805: 2574 (first entry)
 DT 02-APR-1991 (first entry)
 DE Sequence of the BglII/EcoRI DNA fragment of genomic clone 108-1 of
 DE Eimeria tenella encoding the TA antigen.
 KW Poultry vaccine; Eimeria tenella vaccine; Eimeria necatrix vaccine;
 OS Eimeria tenella.
 FR Key Location/Qualifiers
 FT exon 825..912
 FT intron 913..913
 FT /tag= a
 FT /tag= b
 FT /note= "intron A"
 FT 1014..1189
 FT exon 1190..1303
 FT intron 1304..1303
 FT /tag= d
 FT /note= "intron B"
 FT 1304..1665

FT intron /tag= g
 FT 1665.1789 /tag= f
 FT /note= intron C
 FT 1790.1923 /tag= g
 PD 04-JUN-1987
 PD A0865869-A
 PR 01-DEC-1986: 065869
 PR 03-DEC-1985: US-805301.
 PR 04-DEC-1985: US-805301.
 PR 11-DEC-1985: US-805824.
 PR 11-DEC-1985: US-807497.
 PR (SOLV) SOLWAY & CIE.
 PA Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,
 PI Kuhn I, McCaman MT, Files JG, Sias SR.
 PI 81-182
 PI 81-87-199027/23.
 DR P-PSDB: 8170487.
 DR New polypeptide for protecting chickens against coccidiosis - are
 FT obtained by recombinant DNA procedures involving total genomic DNA
 FT from elmeria maxima oocysts.
 FT Claim: 1. Fig 5: 255pp: English.
 CC Elmeria tenella polypeptides are used to immunise chickens against
 CC infection against Elmeria strains, so that coccidiosis is prevented.
 CC They are used in vaccines, e.g. in a dose to provide over 20
 CC microgram/Ag. The monoclonal antibody or its anti-idiotypic antibody
 CC may also be administered to confer protection.
 CC Sequence 2574 BP: 640 A; 594 C; 627 G; 703 T;
 SQ
 Query Match 1.48; Score 16; DB 2; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 Db 222 ggggattattctggtg 237
 QY 159 GCGGATATTCGCTG 174
 RESULT 38
 IT N70804 standard: DNA: 2575 BP.
 AC Q70804:
 DT 02-APR-1991 (first entry)
 DE Sequence of the BglII/EcoRI DNA fragment of genomic clone 108-1 of
 DE Elmeria tenella encoding the TMA antigen.
 KW Polity vaccine; Elmeria tenella vaccine; Elmeria necatrix vaccine;
 KW Elmeria tenella.
 OS Location/Qualifiers
 FT Key 1. 914
 FT /number= 1
 FT /note= "827..829 = translation initiation codon"
 FT /tag= b
 FT /note= intron A
 FT 1016.1191 /tag= c
 FT /number= 2
 FT 1192.1305 /note= intron B
 FT 1306.1667 /tag= e
 FT /number= 3
 FT 1473.1733 /tag= f
 FT /note= intron C
 FT 1792.2424 /tag= g
 FT /note= "1925..1927 = termination codon"
 FT misc_difference 51 /tag= h
 FT /note= "Represents the number 7"
 FT misc_difference 57 /tag= i
 FT /note= "Represents the number 8"
 FT misc_difference 110 /tag= j
 FT /note= "Represents the number 5"
 FT misc_difference 162 /tag= k
 FT /note= "Represents the letter O"
 FT misc_difference /tag= l
 FT /note= "Printed as Y in the specification"
 FT misc_difference 286 /tag= m
 FT /note= "Represents the letter O"
 FT misc_difference 300 /tag= n
 FT /note= "Represents the number 3"

FT intron 913.1013 /tag= b
 FT /note= intron A
 FT 1014.1189 /tag= c
 FT 1150.1303 /tag= d
 FT 1304.1665 /tag= e
 FT 1665.1789 /tag= f
 FT /note= intron
 FT 1790.1923 /tag= g
 PD A0865869-A
 PD 01-DEC-1986: 065869
 PD 03-DEC-1985: US-805301.
 PR 06-DEC-1985: US-805824.
 PR 11-DEC-1985: US-805824.
 PR 11-DEC-1985: US-807497.
 PR (SOLV) SOLWAY & CIE.
 PI Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,
 PI Paul LS, Chang RJ, Andrews WH,
 PI Kuhn I, McCaman MT, Files JG, Sias SR.
 PI 81-182
 PI 81-87-199027/23.
 DR P-PSDB: 8170487.
 DR New polypeptide for protecting chickens against coccidiosis - are
 FT obtained by recombinant DNA procedures involving total genomic DNA
 FT from elmeria maxima oocysts.
 FT Claim: 1. Fig 5: 255pp: English.
 CC Elmeria tenella polypeptides are used to immunise chickens against
 CC infection against Elmeria strains, so that coccidiosis is prevented.
 CC They are used in vaccines, e.g. in a dose to provide over 20
 CC microgram/Ag. The monoclonal antibody or its anti-idiotypic antibody
 CC may also be administered to confer protection.
 CC Sequence 2575 BP: 640 A; 601 C; 624 G; 706 T;
 SQ
 Query Match 1.48; Score 16; DB 2; Length 2575;
 Best Local Similarity 100.0%; Pred. No. 1.86e+01;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 Db 222 ggggattattctggtg 237
 QY 159 GCGGATATTCGCTG 174
 RESULT 39

ID Q37481 standard: DNA: 2577 BP.
 AC Q37481:
 DT 11-JUN-1993 (first entry)
 DE Sequence of the BglII/EcoRI DNA fragment of genomic clone 108-1 of
 DE Elmeria tenella encoding the TMA antigen.
 KW Polity vaccine; Elmeria tenella vaccine; Elmeria necatrix vaccine;
 KW Elmeria tenella.
 OS Location/Qualifiers
 FT Key 1. 914
 FT /number= 1
 FT /note= "827..829 = translation initiation codon"
 FT /tag= b
 FT /note= intron A
 FT 1016.1191 /tag= c
 FT /number= 2
 FT 1192.1305 /note= intron B
 FT 1306.1667 /tag= e
 FT /number= 3
 FT 1473.1733 /tag= f
 FT /note= intron C
 FT 1792.2424 /tag= g
 FT /note= "1925..1927 = termination codon"
 FT misc_difference 51 /tag= h
 FT /note= "Represents the number 7"
 FT misc_difference 57 /tag= i
 FT /note= "Represents the number 8"
 FT misc_difference 110 /tag= j
 FT /note= "Represents the number 5"
 FT misc_difference 162 /tag= k
 FT /note= "Represents the letter O"
 FT misc_difference /tag= l
 FT /note= "Printed as Y in the specification"
 FT misc_difference 286 /tag= m
 FT /note= "Represents the letter O"
 FT misc_difference 300 /tag= n
 FT /note= "Represents the number 3"

FT misc_difference 427 /tag= o
 FT /note= "Printed as Y in the specification"
 FT misc_difference 451 /tag= p
 FT /note= "Printed as M in the specification"
 FT misc_difference 514 /tag= q
 FT /note= "Printed as M in the specification"
 FT misc_difference 615 /tag= r
 FT /note= "Represents the number 7"
 FT misc_difference 673 /tag= s
 FT /note= "Represents the number 7"
 FT misc_difference 1778 /tag= t
 FT /note= "Represents the number 7"
 FT 05187080-A
 FT 05-JUN-1984: 617483.
 PR 05-JUN-1984: US-617483.
 PR 16-MAY-1985: US-734085.
 PR 06-DEC-1985: US-805824.
 PR 11-DEC-1985: US-807497.
 PR (SOLV) SOLWAY & CIE.
 PA Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,
 PI Kuhn I, McCaman MT, Files JG, Sias SR, Tedesco JL,
 PI 81-182
 PI 81-87-199027/23.
 DR P-PSDB: 8170487.
 DR Nucleic acid encoding Elmeria tenella proteins - capable of
 FT inducing immune response in chickens, useful in vaccine
 FT Example 6: Fig 5: 255pp: English; DNA isolated from Elmeria
 FT tenella oocysts.
 CC The library was screened with oligonucleotide
 CC probes based on peptide fragments of the purified 17,000 Dalton
 CC component of the TMA antigen. Clone 108-1 hybridised strongly to
 CC one probe and weakly to the other. This positive clone was
 CC subcloned and sequenced. The sequence of the TMA antigen
 CC from the clone insert was sequenced. It encoded a 23 amino acid
 CC signal peptide which was removed from the mature, purified protein.
 CC The DNA also contained 3 introns. From the predicted amino acid
 CC sequence of the TMA antigen (17 kD) it seems that the two
 CC contiguous nucleotide sequences, and at least one proteolytic step
 CC occurs to generate the 8 kD peptide by removal of a tripeptide.
 CC The sequence as printed in the specification includes several
 CC errors (indicated by asterisks) which have been replaced by N
 CC for inclusion in the Genbank record.
 CC Sequence 2577 BP: 640 A; 594 C; 627 G; 703 T;
 SQ

[illegible]

CC	sequence provides high detection sensitivity and specificity. The
CC	probe may be optionally labelled with biotin and used in methods to
CC	quantitate the amount of DNA in a sample. See 62732-51.
SC	Sequence 40 BP; 12 A; 7 C; 14 T;
	Query Match 1.3%; Seed 14; DB 11; Length 40;
	Best Local Similarity 100.0%; Pref.No. 3.47e+02;
	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	12 tgcgaactacctttt 25
OY	
	941 TGCAAGAACTACTTTT 954
RESULT	59
ID	T07129 standard; cDNA; 48 BP.
AC	T07129; 89% (first entry)
KW	Human LAP 48-mer probe
DN	Lingual antimicrobial peptide; epithelium; LAP; bovine; beta-defensin;
KM	antimicrobial activity; Gram-positive bacteria; Gram-negative bacteria;
KV	fungal pathogens; mammal; microbial infection; immunodeficiency; AIDS;
PF	Synthetic; gum disease; burn; pneumonia; human; probe; ss.
PH	W09532287-A1.
PC	30-MAY-1995. D06761
PD	24-MAY-1994; US-248016
PI	(MAGA.) MAGANIN PHARM INC.
P1	Schönwetter BS, Zasloff MK:
PT	New active 05B27/02 peptide from mammalian lingual epithelium - for
PT	treating bacterial or fungal infections, esp. of the epithelium,
PT	e.g. in gum disease, cystic fibrosis, burns, etc.
PS	[disclosure] Page 5; 37pp. English.
CC	This sequence represents the LAP
CC	lingual antimicrobial peptide (LAP). This sequence represents the LAP
CC	48-mer probe. LAP is a member of the beta-defensin group of the defensins
CC	family of peptides. LAP has broad spectrum antimicrobial activity
CC	against Gram-positive and Gram-negative bacteria, and fungal pathogens
CC	and levels induced in response to epithelial injury/infection. The
CC	cDNA encoding bovine LAP (see T07133) can be used in a method of
CC	identifying endogenous up-regulators of LAP. In this method, epithelial
CC	cells are treated with various substances and the presence of LAP mRNA
CC	is then measured to determine whether the substance is an
CC	up-regulator. LAP is used to treat microbial infections of the
CC	epithelium (or of those that extend to deeper tissues) e.g. in
CC	bummed client states (AIDS), cystic fibrosis, gum disease, wounds,
SC	Sequence 48 BP; 8 A; 15 C; 11 G; 14 T;

Tue Nov 17 08:55:26 1998

US-08-887-977-9. rmd

EP-594959-A.
04-MAY-1993. 113261.
19-AUG-1993. 113261.
28-AUG-1993. US-93084.
P1 REYNOLDS R.L., WALSH P.S.,
REYNOLDS R.L., WALSH P.S.,
WPI; 94-18573/18.
DR Oligonucleotide probes for detecting human DNA - having
PT conserved complementary to the human alpha satellite locus or a
PT conserved complementary to the human mitochondrial genome
PT Claim 6, page 17; 27pp; English.
PC This oligonucleotide probe is complementary to a region in the 2.7
CC kilobase D17L1 locus, a highly repetitive sequence. The use of
CC this probe allows for high detection sensitivity and specificity. The
CC probe may be optionally labelled with biotin and used in methods to
CC quantitate the amount of DNA in a sample. See 02137-51.
SQ Sequence 80 BP; 24 A; 15 C; 16 G;
Query Match 1.3%; Score 14; DB 11; Length 80;
Beat Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 tcgaactactctt 25
|||||
Gy 941 TCGAACACTACTT 954

RESULT 65
ID T4568 standard; cDNA to mRNA; 80 BP.
DC 75-SEP-1996 (first entry)
DE Human gene signature HMGSO6618.
DW Gene signature; messenger RNA; mRNA relative abundance; frequency;
KW Human cloning; mapping; non-biased library; diagnosis; detection;
OS Homo sapiens. Abnormal cell function; ss.
PN W0951472.A1.
PD 01-JUN-1995. 301916
PE 12-NOV-1993. JP-355504.
PF 12-NOV-1993. JP-355504.
PG (MATSU)/ MATSUBARA K.
PA (OKUBU)/ OKUBO K.
PP Watanabe, Nobuo K.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
CC Class 1, Page 1641. 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.

[illegible]

The Nov 17 08:55:26 1998

Tue Nov 17 08:55:26 1998

Another specific cDNA or genomic sequences can be used to identify and isolate another specific promoters. The another specific promoter can then be cloned into a recombinant construct and used to express heterologous genes. A preferred heterologous genes include Diphtheria toxin A-chain gene, Pertussis toxin gene from *Bordetella pertussis*, the *gus* gene from SV40, the *lacZ* gene from phage M13, the indole acetic acid-lyase synthetase gene from *Pseudomonas syringae* and the OYA toxin gene from *Bacillus thuringiensis israelensis*. All of these genes when transformed into tobacco protoplasts will produce transgenic plants which plant to produce viable pollen. Transformation of plants with such a recombinant construct can produce transgenic, male sterile plants. Male sterility is important in the production of hybrid seeds as it prevents self pollination which hinders breeding and hybrid seed sequence 95 bp; 34 C; 24 G; 12 G; 25 T;

Query Match 13 A; Score 14; DB 10; Length 95;
Mismatches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches 14

DB 21 attcttaaaaggca 34
QY 708 ATTCTTAAGGCA 722

RESULT 67 standard; DNA, 100 BP.
AC AGCAGTGGTTATCGA
AD GGTGATG
AG GGTGATG
AG GGTGATG
AG 06-DEC-1994 (first entry)
DE oligonucleotide probe for repeated sequence of D1B1 locus.
EY oligonucleotide probe for repeated sequence of D1B1 locus.
FM conserved sequence, human alpha satellite locus, mitochondrial, ss.
GN Synthetic
PN EP-594959-A
PR 19-MAY-1993, 113261
PR 19-MAY-1993, 113261
PR 28-AUG-1992; US-931084.
PR REYNOLDS KL, Walsh PS:
PA REYNOLDS KL, Walsh PS:
PI (HOFF) HOFFMANN LA ROCHE & CO AG F.
PT oligonucleotide probes for detecting human DNA - having
PT sequences complementary to a human alpha satellite locus or a
PT conserved sequence in the mitochondrial genome
PT Taim G, George J, 27pp English
PT 16-SEP-1993, 113261
CC kilobase D1B1 locus, a highly repetitive sequence to a region in the 2.7
CC probes complementary to a highly repeated sequence or to a conserved
CC sequence provides high detection sensitivity and specificity. The
CC quantitates the amount of DNA in a sample. See 062331-31
CC

```

6Q Sequence 100 BP: 30 A; 18 C; 20 G; 32 T;
Query Match 1.3%; Score 14; DB 11; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 tgaagcaactctt 25
   ||| ||||| |||||
Qy 941 TCAGAACTACTT 954

RESULT 68
ID T22631 standard; cDNA to mRNA; 110 BP.
DT 07-DEC-1996 (first entry)
DE Human gene signature HMG503934.
DK Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW Kwon; cloning; mapping; non-biased library; diagnosis; detection;
KW Kwon; cloning; mapping; non-biased library; diagnosis; detection;
KW Homo sapiens; abnormal cell function; as.
KW Homo sapiens
PN PM W0951472-AL.
PD 01-JUN-1995. 301915
PD 01-JUN-1995. 301915
PD 12-NOV-1993. JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKEBU) OKEBU K.
PI Matsubara K.; Okebu K.
PT Identifying gene signatures in 3'-directed human cDNA library. A e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissue.
PI Claim 1, Page 1107; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T7900078837 and which is able to hybridise to part of
CC the 7837 "GS" sequences, and a library of cDNA libraries prepared
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(t) as the sole primer. Since the 3'-
CC end of the corresponding cDNA hybridized to a particular "GS" sequence, almost
CC all the 3'-end cDNA hybridized to a particular "GS" sequence, almost
CC all the 3'-end cDNA hybridized to a particular "GS" sequence, almost
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined as a means of diagnosing abnormal cell function or for
CC recognizing different cell types.
CC sequence 110 BP; 32 C; 22 G; 12 A; 37 T;
Query Match 1.3%; Score 14; DB 20; Length 110;

```

12-125-1993; 001294.

Page 97

THE

Page 91

CC This sequence represents the cDNA encoding bovine epithelial lingual antimicrobial peptide (LAP). LAP is a member of the beta-defensin group of peptides that have been shown to have antimicrobial activity against Gram-positive and Gram-negative bacteria, and fungal pathogens. LAP is present at low levels in mammalian epithelia, with high expression levels being induced in response to epithelial injury/infection. This sequence can be used in a method of identifying and isolating genes that encode antimicrobial peptides. The LAP gene was cultured in the presence of a test substance. The levels of LAP mRNA are then measured to determine whether the substance is an up-regulator. LAP is used to treat microbial infections of the epithelium (or of those that are used to deeper tissues), e.g. in immunodeficient states (AIDS), cystic fibrosis, 127 BP; 39 A; 26 C; 38 G; 24 T;

Query Match 1.39; Score 15; DB 18; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 90 cgtctctcgagccca 104
Q7 581 CAGCTCTCGAGCCCA 595

RESULT 74

ID 065335 standard; DNA; 130 BP.
AC 207114; 1996 (first entry)
DE Oligonucleotide probe for repeated sequence of D1721 locus.
KW Probe: detection; repeat sequence; D1721; locus; control region;
OS Contaminant; human alpha satellite locus; mitochondria; ss.
PA EP-594959-A.
PD 04-MAY-1994.
PF 15-AUG-1993; 113261.
PI (HOPF) ROSEFARN LA ROCHE & CO AG F.
P1 Reynolds RL, Walsh PS:
P2 WPI: 94-145673/18.
P3 This oligonucleotide probe for detecting human DNA - having conserved sequence in the mitochondrial genome.
P4 Claim 6; Page 14; 27pp; English.

CC This oligonucleotide probe is complementary to a region in the 2.7 kbase D1721 locus, a highly repetitive sequence. The use of this probe provides high detection sensitivity and specificity. The probe may be optionally labelled with biotin and used in methods to quantitate the amount of DNA in a sample. See 062532-51.
CC Sequence 130 BP; 37 A; 23 C; 30 G; 40 T;
CC

SQ Sequence 133 BP; 40 A; 27 C; 38 G; 28 T;
Query Match 1.39; Score 15; DB 18; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 96 cgtctctcgagccca 110
Q7 581 CAGCTCTCGAGCCCA 595

RESULT 76

ID 031133 standard; DNA; 152 BP.
AC 212523; 1993 (first entry)
DE ICAM-3 gene fragment #2.
KW Interleukin adhesion molecule-3; ICAM-3; ICAM-1; ICAM-2; homology; NK-17; Ig domain; NK-10; domain 4; probe; leukocyte; lymphocyte;
OS Homo sapiens; neutrophil; inflammation; immune response; ss.
PA Key Key Location/Qualifiers
PF cds 2..151
P1 W09222323-A.
P2 21-DEC-1992.
P3 11-JUN-1992; U04896.
P4 11-JUN-1992; US-712879.
P5 11-JUN-1992; US-712879.
P6 11-JUN-1992; US-712879.
P7 11-JUN-1992; US-712879.
P8 WPI: 93-017508/02.
P9 P-PSDB; R30475.
P10 Inter-cellular adhesion molecule-3 inhibiting granulocyte, lymphocyte, monocyte, and macrophage for cell adhesion, etc.
P11 asthma, auto-immune thyroiditis, multiple sclerosis, AIDS
P12 Claim 47; Page 79; 173pp; English.

CC The sequences given in Q31132-14 are fragments of the interleukin adhesion molecule-3 (ICAM-3) gene. The sequences are used as probes by which populations of leukocytes recognize and adhere to cellular substrates. ICAM-3 mediates cellular interactions with other lymphocytes, macrophages and neutrophils at the site of inflammation and sites of immune responses. 38 C; 50 G; 38 T;
CC Sequence 132 BP; 28 A; 38 C; 50 G; 38 T;
CC

Query Match 1.39; Score 14; DB 6; Length 152;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 28 gaacccctgctct 11
Q7 906 GAAACCCCTGCTCT 919

Query Match 1.39; Score 14; DB 11; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 12 tcgaactactctt 25
Q7 941 TCGAACACTACTCTT 954

RESULT 75

ID 707134 standard; cDNA; 133 BP.
AC 207114; 1996 (first entry)
DE Pro-LAP coding sequence.
KW Lingual antimicrobial peptide; epithelium; LAP; bovine; beta-defensin; antimicrobial activity; Gram-positive bacteria; Gram-negative bacteria; fungal pathogen; mammal; microbial infection; immunodeficiency; AIDS;
OS Homo sapiens; gum disease; burn; pneumonia; ss.
PA Bos taurus. Location/Qualifiers
PF Key signal_peptide 84..89
P1 W0951297-A1.
P2 24-MAY-1994; 006761.
P3 24-MAY-1994; US-248016.
P4 (MAGA-) MAGALININ PHARM INC.
P5 Schoneveldt JBS, Fasloff WA;
P6 WPI: 93-017508/02.
P7 P-PSDB; R66895.

CC New antimicrobial peptide from mammalian lingual epithelium - for treating bacterial or fungal infections, esp. of the epithelium, C-9 in gum disease, cystic fibrosis, burns, etc.
CC This sequence represents the cDNA encoding the pro form of bovine epithelial lingual antimicrobial peptide (LAP). LAP is a member of the beta-defensin group of the defensin family of peptides. LAP has broad spectrum antimicrobial activity against Gram-positive and Gram-negative bacterial and fungal pathogens. LAP is induced in response to epithelial injury/infection. This sequence can be used in a method of identifying endogenous up-regulators of LAP. In this method, the levels of LAP mRNA are then measured to determine whether the substance is an up-regulator. LAP is used to treat microbial infections of the epithelium (or of those that extend to deeper tissues) e.g. in immunodeficient states (AIDS), cystic fibrosis, gum disease, wounds, burns and pneumonia.

RESULT 77

ID 031626 standard; DNA; 164 BP.
AC 02-FEB-1993 (first entry)
DE Upstream sequence of microsatellite from clone GBKCA5.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
OS Genetic mapping; traits; amplification; ss.
PA W09213102-A.
PD 06-APR-1992.
PF 15-JUN-1992; U00340.
P1 15-JUN-1992; US-642342.
P2 (GENE-) GENMARK.
P3 Georges M, Massey JM;
P4 WPI: 92-284684/34.
P5 Polymorphic bovine DNA markers - used in genetic identification, breeding and selection studies.
P6 Table 7; Page 180; 51pp; English.

CC The sequence is that upstream of a bovine microsatellite sequence obtd. by screening a library of bovine MboI DNA fragments of between 350 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe. CC Oligonucleotide primers were used to amplify the sequence. CC The distribution of microsatellites and MboI sites the frequency of (76)n > 9 microsatellites in the bovine genome is estimated at >100, CC 000. The sequence information for ca. 230 such bovine microsatellites was submitted in the specification and linked herein (see below). CC The sequences of the microsatellites are given in the table. CC These sequences were used to generate the required PCR primers for in vitro amplification of the corresp. microsatellite (using the program OPTIPRIM). The microsatellites may be used to identify individuals, CC loci, or genes involved in the determination of economically important traits esp. in cattle, to allow selective breeding.
CC See also G33501-34437.
CC Sequence 164 BP; 63 A; 35 C; 24 G; 43 T;
CC

Query Match 1.39; Score 14; DB 5; Length 164;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 71 aacacacacaccta 84
Cp 756 AAGCACCACACCTA 743

RESULT 78

ID Q97185 standard; DNA; 167 BP.
AC Q97185;
DT 01-APR-1996 (first entry)

Human type I steroid 5-alpha reductase cDNA recombinant fragment.
 SW 5-alpha reductase cDNA recombinant fragment.
 KW Probe; recombinant; inhibits prostatic hyperplasia; gene; hirsutism;
 KW male pattern baldness; endometriosis; prostate cancer; testosterone;
 KW dihydroxytestosterone; es.
 PN US542262-A.
 PD 06-JUN-1995.
 PF 30-APR-1990; 517661.
 FR 30-APR-1990; US-517661.
 PA (TCEA) UNIV TEXAS SYSTEM.
 PI Anderson S, Russell DW.
 DR Steroid 5 alpha-reductase nucleic acid segments and recombinant
 PT Steroid 5 alpha-reductase nucleic acid segments and recombinant
 PT and abnormal sexual differentiation.
 PS Claim 16: Column 61-62; 72pp; English.
 CC 5-alpha reductase enzymes catalyze the conversion of testosterone to
 CC dihydroxytestosterone. The human steroid 5-alpha reductase I (SRD5A-I)
 CC cDNA has been cloned and expressed in human enzyme cDNA or fragments of
 CC this (C97482-097394) can be used in the human enzyme cDNA or fragments of
 CC for the large scale production of SRD5A or as probes for enzyme-encoding
 CC sequences from alternative sources. The sequences are also useful in
 CC the analysis of normal and abnormal sexual differentiation, benign
 CC endometriosis, prostatic hyperplasia, male pattern baldness; gene; hirsutism and
 CC endometriosis.
 SQ Sequence 167 BP; 40 A; 33 C; 42 G; 52 T;

Query Match 1.38; Score 14; DB 17; Length 167;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 56 ttcataactggcat 69
 Cp 419 ttcataactggcat 406

RESULT 79
 ID T25902 standard; cDNA to mRNA; 175 BP.
 AC W1324.
 AD W1324.
 DE Probe (sub 1) to determine human individual genetic identity.
 DE Probe (sub 1) to determine human individual genetic identity.
 DE DNA probe; HeLa cell; HL-60; K-562; polymorphism;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 OS Homo sapiens.
 PN J6219530C-A.
 PD 28-AUG-1987.
 PF 19-FEB-1986; 034146.
 FR 19-FEB-1986; JP-034146.
 PA (NICH) NICHIREI KK.

DR WPI: 87-280952.
 PT DNA probe for assaying individual genetic identity - obtained from
 PT HeLa cell DNA.
 PS Claim 1, page 586; 6pp; Japanese.
 CC The sequence is obtd. from HeLa cell DNA, HL-60 cell DNA, or K-562
 CC cell DNA and is used for individual gene identity assay, by method
 CC of hybridization with polymorphic DNA in human genome.
 CC See also N70876 and W1324-28.
 SQ Sequence 171 BP; 55 A; 31 C; 35 G; 50 T;

Query Match 1.38; Score 14; DB 1; Length 171;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 52 tcagaactacttt 65
 Cp 941 TCAGAACTACTTT 954

RESULT 80
 ID W1328 standard; DNA; 171 BP.
 AC W1328.
 AD W1328.
 DE Probe (sub 5) to determine human individual genetic identity.
 DE DNA probe; HeLa cell; HL-60; K-562; polymorphism;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 OS Homo sapiens.
 PN J6219530C-A.
 PD 28-AUG-1987.
 PF 19-FEB-1986; 034146.
 FR 19-FEB-1986; JP-034146.
 PA (NICH) NICHIREI KK.

Query Match 1.38; Score 14; DB 1; Length 171;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 52 tcagaactacttt 65
 Cp 941 TCAGAACTACTTT 954

RESULT 81
 ID T25902 standard; cDNA to mRNA; 175 BP.
 AC W1324.
 AD W1324.
 DE Human gene signature HNGS00739.
 DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 OS Homo sapiens.
 PN J6219530C-A.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 FR 12-NOV-1993; JP-355504.
 PA (OKUB) OKUBO K.
 PI Matabara K, Okubo K.
 DR WPI: 95-206931/27.
 DE Identifying gene signatures in 3'-directed human cDNA library - e.g.
 DE for diagnosis of abnormal cell function by preparing cDNA that
 DE reflects relative abundance of corresp. mRNA in specific human
 PT tissues.
 PS Claim 1, page 195; 224pp; Japanese.
 CC Double-stranded DNA (or its complementary strand or the corresp.
 CC given in T19001-726837) which is able to hybridize to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from human tissues. The synthesis of cDNA was initiated from
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC the corresponding mRNAs in the particular tissue from which it was
 CC derived.
 CC The appearance frequency of given GS in cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognizing different cell types.
 SQ Sequence 175 BP; 51 A; 21 C; 27 G; 67 T;

Query Match 1.38; Score 14; DB 22; Length 175;
 Best Local Similarity 93.3%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 143 tctcttttaagaa 157
 Cp 186 TCTCTTTTAAGAA 200

RESULT 82
 ID T19669 standard; cDNA to mRNA; 179 BP.
 AC T19669.

Query Match 1.38; Score 14; DB 18; Length 179;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 8 ccaagggtttgaca 21
 Cp 697 CCAGGTTTGACA 684

RESULT 83
 ID G61339 standard; DNA; 216 BP.
 AC G61339.
 AD G61339.
 DE Human brain expressed sequence tag EST01339.
 DE Human brain expressed sequence tag EST01339.
 KW Gene transcription product; genetic markers; tagging; in vivo;

PA (MATSU) MATSUBARA K.
 PI (OUSA) MATSUBARA K.
 PI Matsubara K. Okubo K.
 DR WPI: 95-206931/27.
 PR Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function. By preparing cDNA that
 PT tissues relative abundance of corresp. mRNA in specific human
 PT tissues.
 PS Claim 1: Page 1378: 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC double-stranded DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(I) as the sole primer. Since the 3'-end
 CC of mRNA is known, the GS sequences can be determined. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The GS sequences are then used as probes in a cDNA library, can be
 CC determined (e.g. by Northern blotting) and the relative abundance of
 CC sequences as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 307 BP: 101 A; 45 C; 60 G; 99 T;
 Query Match 1.38; Score 14; DB 20; Length 307;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 166 ttcattgattatt 179
 QY 658 ttcattgattatt 671

RESULT 97
 ID G61112 standard: DNA: 311 BP.
 AC G61112;
 DT 16-MAR-1994 (first entry)
 KW Human Brain Expressed Sequence Tag EST0155.
 KW transcribed; mapping; locations: chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN W09116178-A.
 PP 12-FEB-1993: U01294.
 PP 12-FEB-1993: US-837195.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno NP, Venter CJ.
 PT Enriched oligonucleotides and corresp. sequences - used as

PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC to specific chromosomes, for identifying locations of genes in
 CC locations of disease-associated genes, for identification of trans-
 CC type, and for prepn. of antisense sequences, probes and constructs.
 CC EST01155 has a "Poor" coding probability as evaluated using the
 CC coding-region prediction program CDM. See also 059041-061440.
 SQ Sequence 311 BP: 30 A; 60 C; 56 G; 92 T;
 Query Match 1.38; Score 14; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 49 acccttcctttta 62
 QY 181 acccttcctttta 194

RESULT 98
 ID Q7289 standard: DNA: 327 BP.
 AC Q7289.
 DT 1994 (first entry)
 KW Human genome fragment (Preferred).
 KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
 KW detection; homology; human; adrenal tissue; da.
 OS Homo sapiens.
 PN W09116178-A.
 PP 20-JAN-1994.
 PP 13-JUL-1993: G01467.
 PR (MEDR) MEDICAL RES COUNCIL.
 PI (MEDI) MEDICAL RES COUNCIL.
 PI Silson DR, Starkey M.
 DR WPI: 94-015056/04.
 PT New nucleic acid fragment encoding gene products - can be used
 PT for genetic analysis and mapping.
 CC Human nucleic acid fragments; isolated from brain, adrenal tissue.
 CC the placenta or bone marrow comprise any of: (A) a sequence
 CC selected from (076401-077613), (B) an allelic variation of a
 CC sequence as described in (A), or (C) a sequence complementary
 CC to sequence as described in (A), or (C) a sequence complementary
 CC Preferred sequences exhibit no more than 90% homology to a human
 CC sequence known per se.
 SQ Sequence 327 BP: 122 A; 42 C; 52 G; 111 T;

Query Match 1.38; Score 14; DB 11; Length 327;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 242 gttacacacat 255
 QY 677 gttacacacat 684

RESULT 99
 ID T24692 standard: cDNA to mRNA: 330 BP.
 AC T24692;
 DT 09-OCT-1996 (first entry)
 KW Human Brain Expressed Sequence Tag EST04760.
 KW Gene signature; messenger RNA cDNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09116178-A.
 PP 01-JUN-1998.
 PP 12-NOV-1994: J01916.
 PR 12-NOV-1993: JP-35504.
 PA (MATSU) MATSUBARA K.
 PI Matsubara K. Okubo K.
 DR WPI: 95-206931/27.
 PR Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function. By preparing cDNA that
 PT tissues relative abundance of corresp. mRNA in specific human
 PT tissues.
 PS Claim 1: Page 1671: 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC double-stranded DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(I) as the sole primer. Since the 3'-end
 CC of mRNA is known, the GS sequences can be determined. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The GS sequences are then used as probes in a cDNA library, can be
 CC determined (e.g. by Northern blotting) and the relative abundance of
 CC sequences as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 330 BP: 97 A; 61 C; 70 G; 99 T;
 Query Match 1.38; Score 14; DB 21; Length 330;
 Best Local Similarity 93.3%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 264 tttataaagcagaa 278
 QY 197 tttataaagcagaa 183

RESULT 100
 ID Q43746 standard: cDNA: 330 BP.
 AC Q43746.
 DT 1993 (first entry)
 KW Human Brain Expressed Sequence Tag EST04760.
 KW Gene signature; messenger RNA cDNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09116178-A.
 PP 01-JUN-1998.
 PP 12-NOV-1994: J01916.
 PR 12-NOV-1993: JP-35504.
 PA (MATSU) MATSUBARA K.
 PI Matsubara K. Okubo K.
 DR WPI: 95-206931/27.
 PR Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function. By preparing cDNA that
 PT tissues relative abundance of corresp. mRNA in specific human
 PT tissues.
 PS Claim 1: Page 1671: 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC double-stranded DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(I) as the sole primer. Since the 3'-end
 CC of mRNA is known, the GS sequences can be determined. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The GS sequences are then used as probes in a cDNA library, can be
 CC determined (e.g. by Northern blotting) and the relative abundance of
 CC sequences as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 330 BP: 97 A; 61 C; 70 G; 99 T;
 Query Match 1.38; Score 14; DB 21; Length 330;
 Best Local Similarity 93.3%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 1.34; Score 14; DB 7; Length 330;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 251 agctggagcttg 264
 |||||||
 Cp 1009 ACCCGAGGCTTG 396

RESULT 101
 ID T27968 standard; DNA: 335 BP.
 AC Q77282;
 DT 23-SEP-1994 (first entry)
 DE Human genome fragment (Preferred).
 KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;
 RN Relative abundance; frequency; diagnosis; detection;
 KW Cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W091472-A1.
 PD 11-JUN-1995; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATSU) MATSUBARA K.
 PA (ORUBU) ORUBO K.
 DB WPI: 95-206317/27; Pubo K.
 WP: 95-206317/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 cell types.
 PS Claim 1. Page 931: 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences
 CC given in T27968 and which is able to hybridize to part of
 CC the 7837 "Gs" sequences, is used for diagnosing abnormal cell function
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence unique to a particular mRNA species, almost
 CC all the 3'-end of cDNA molecules obtained from the 3'-directed library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined by comparing the appearance frequency of the GS
 CC sequences obtained from the 3'-directed library with the appearance
 CC frequency of different cell types.
 CC Recognizing different cell types.
 CC Sequence 335 BP; 108 A; 62 C; 41 G; 120 T;

Query Match 1.34; Score 14; DB 19; Length 335;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

cell typing; abnormal cell function; ss.

OS Homo sapiens.
 PN W091472-A1.
 PD 11-JUN-1995; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATSU) MATSUBARA K.
 PA (ORUBU) ORUBO K.
 DB WPI: 95-206317/27; Pubo K.
 WP: 95-206317/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 cell types.
 PS Claim 1. Page 2095: 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences
 CC given in T27968 and which is able to hybridize to part of
 CC the 7837 "Gs" sequences, is used for diagnosing abnormal cell function
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence unique to a particular mRNA species, almost
 CC all the 3'-end of cDNA molecules obtained from the 3'-directed library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined by comparing the appearance frequency of the GS
 CC sequences obtained from the 3'-directed library with the appearance
 CC frequency of different cell types.
 CC Recognizing different cell types.
 CC Sequence 340 BP; 61 A; 62 G; 114 T;

Query Match 1.34; Score 14; DB 21; Length 340;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 267 cttttgtttttt 280
 |||||||
 Qy 631 CTTTGTGTTTCT 644

RESULT 104
 ID T27968 standard; DNA: 340 BP.
 AC Q31044;
 DT 24-MAR-1993 (first entry)
 DE HCV-1 genotype GII NSS region sequence na59h6.
 KW Hepatitis C virus; non-A, non-B hepatitis; NSS region; ss.
 RN Relative abundance; frequency; diagnosis; detection;
 PN W09219743-A.
 PD 12-NOV-1992.
 PR 08-MAY-1992; 004036.

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 ttgtgttttttt 45
 |||||||
 Qy 184 TTGCTTTTATTA 197

RESULT 102
 ID Q77282 standard; DNA: 337 BP.
 AC Q77282;
 DT 23-SEP-1994 (first entry)
 DE Human genome fragment (Preferred).
 KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;
 RN Relative abundance; frequency; diagnosis; detection;
 KW Cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09401548-A.
 PD 13-JUN-1993; G01467.
 PR 13-JUN-1993; G01467.
 PA (MEDI-) MEDICAL RES COUNCIL.
 PA Gross J, Radfield M, Howells D, Kelly M, Shaw D;
 PI Shewell M, Howells D;
 DB WPI: 94-015056/04.
 WP: 94-015056/04.
 PT New nucleic acid fragment encoding gene products - can be used
 PT for genetic analysis and mapping
 PS Claim 1. Page 434-435: 616pp; English; from brain, adrenal tissue,
 PS the placenta or bone marrow comprise any of: (A) a sequence
 CC selected from (Q76401-Q77613), (B) an allelic variation of a
 CC sequence as described in (A), or (C) a sequence complementary
 CC to (A) or (B) sequences exhibit no more than 90% homology to a human
 CC sequence known per se.
 CC Sequence 337 BP; 88 A; 83 C; 55 G; 111 T;

Query Match 1.34; Score 14; DB 11; Length 337;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 actgagtaattaga 52
 |||||||
 Cp 58 ACTGAGTATATAGA 45

RESULT 103
 ID T26475 standard; cDNA to mRNA, 340 BP.
 AC T26475;
 DT 23-OCT-1986 (first entry)
 DE Human gene signature HMG508717.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 RN Relative abundance; frequency; diagnosis; detection;
 KW Human, cloning; mapping; non-blended library; diagnosis; detection;

08-MAY-1991; US-697326.
 PR (CHIR) CHIRON CORP.
 PA Beall E, Cha T, Irvine B, Kolberg J, Urdea MS;
 PI Kolberg J, Urdea MS;
 CC Comparing a non-hepatitis C virus-1 nucleotide sequence
 PT - related to HCV-1, useful for treating and detecting HCV-1
 PT infections and as a vaccine
 PS Claim 4: Page 62; 186pp; English.
 CC The sequence within the NSS region of HCV-1 of genotype GII 1b
 CC is a preferred example of an oligonucleotide for use as a probe in
 CC hybridization assays, as a primer for synthesis of HCV genotype-
 CC specific nucleic acid, as a binding partner for separating HCV
 CC expression of HCV genes. Polypeptides encoded by oligonucleotides of
 CC the invention (no sequences given in the specification) are useful
 CC in vaccines against HCV and to produce antibodies to detect the
 CC virus.
 CC Sequence 340 BP; 75 A; 97 C; 104 G; 64 T;

Query Match 1.34; Score 15; DB 5; Length 340;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttctgtctatctgc 299
 |||||||
 Cp 1076 TTCTGTCTATCTGC 1062

RESULT 105
 ID T27968 standard; DNA: 340 BP.
 AC T27968; 1997 (first entry)
 DE Hepatitis C virus type 1 isolate PR2 bases 7912-8771.
 KW Hepatitis C virus; subtype; polymerase chain reaction; amplification;
 RN PCR; primer; probe; antibody; infection; ss.
 OS Hepatitis C virus.
 PN W09613590-A1.
 PD 09-MAY-1996.
 PR 23-OCT-1995; 204155.
 PR 21-OCT-1994; EP-870166.
 PR 28-JUN-1995; EP-870076.
 PR (CHIR) CHIRON CORP.
 PI Maertens G, Koster G, Stuyver L;
 DB WPI: 96-251460/25.
 WP: 96-251460/25.
 P-PDB: R96557.

Query Match 1.34; Score 15; DB 5; Length 340;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttctgtctatctgc 299
 |||||||
 Cp 1076 TTCTGTCTATCTGC 1062

RESULT 105
 ID T27968 standard; DNA: 340 BP.
 AC T27968; 1997 (first entry)
 DE Hepatitis C virus type 1 isolate PR2 bases 7912-8771.
 KW Hepatitis C virus; subtype; polymerase chain reaction; amplification;
 RN PCR; primer; probe; antibody; infection; ss.
 OS Hepatitis C virus.
 PN W09613590-A1.
 PD 09-MAY-1996.
 PR 23-OCT-1995; 204155.
 PR 21-OCT-1994; EP-870166.
 PR 28-JUN-1995; EP-870076.
 PR (CHIR) CHIRON CORP.
 PI Maertens G, Koster G, Stuyver L;
 DB WPI: 96-251460/25.
 WP: 96-251460/25.
 P-PDB: R96557.

CC response to epithelial injury/infection. This sequence can be used in a
CC method of identifying endogenous up-regulators of LAP. In this method,
CC levels of LAP mRNA are then measured to determine whether the substance
CC is an up-regulator. LAP is used to treat microbial infections of the
CC epithelium (or of those that extend to deeper tissues) e.g. in
CC immunodeficient states (AIDS), cystic fibrosis, gum disease, wounds,
CC and other conditions. 120 A; 72 C; 85 G; 73 T;
SQ Sequence 350 BP; 1.3% Score 15; DB 18; Length 350;

Query Match 1.3% Score 15; DB 18; Length 350;
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 179 cgtctctggagccca 193

Qy 581 cttctctggagccca 595

RESULT 110

ID Q55233 standard; DNA: 354 BP.
DE 14-JUN-1994 (first entry)
DE Grapevine ribosomal clone 176A.
RW Ribosome; grapevine; vitis; IGS region; rDNA; polymorphism;
RW Key: cultivar; probe; primer; detection; ds.
OS Vitis vinifera. Location/Qualifiers
FH Key primer_bind 202..224
FT primer_bind /*tag= a
FT primer_bind /*tag= b
FN M9401580-A.
PD 20-JAN-1994.
PF 30-JUN-1993; AUG320.
PI (GSR) COMONWELTH SCI & IND RES ORG.
PI Scott NS, Thomas ME;
DR WPI: 94-035083/04.
RT Revi ribosome DNA probe sequences - for the accurate identification
P6 Claim 26: Fig 5c: 55pp; English.
CC The sequences given in Q55231-50 are derived from ribosomes of the
CC grapevine genus vitis. These sequences represent the IGS region of
CC the ribosomal rDNA repeat and contain polymorphisms. These
CC sequences are derived from different grape cultivars. In these clones
CC sequences and were identified in a genomic library of grapevine DNA
CC using simple di-, tri- or tetra- nucleotide repeats such as (AT)₈.
CC (GT)₁₀, (CT)₁₀ and such like as probes. 70 G; 92 T;
SQ Sequence 354 BP; 1.1% Score 14; DB 18; Length 354;

Query Match 1.3% Score 15; DB 18; Length 354;
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 328 ttcattactactcagt 342

Qy 44 ttcattactactcagt 58

RESULT 111

ID T22890 standard; cDNA to mRNA. 358 BP.
DE 29-AUG-1996 (first entry)
DE Human gene signature HMG504614.
RW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
RW Human cloning; mapping; non-biased library; diagnosis; detection;
OS Homo sapiens. Location/Qualifiers
FH Key primer_bind 202..224
FT primer_bind /*tag= a
FT primer_bind /*tag= b
FN M9401580-A.
PD 20-JAN-1994.
PF 30-JUN-1993; AUG320.
PI (GSR) COMONWELTH SCI & IND RES ORG.
PI Scott NS, Thomas ME;
DR WPI: 94-035083/04.
RT Revi ribosome DNA probe sequences - for the accurate identification
P6 Claim 26: Fig 5c: 55pp; English.
CC The sequences given in Q55231-50 are derived from ribosomes of the
CC grapevine genus vitis. These sequences represent the IGS region of
CC the ribosomal rDNA repeat and contain polymorphisms. These
CC sequences are derived from different grape cultivars. In these clones
CC sequences and were identified in a genomic library of grapevine DNA
CC using simple di-, tri- or tetra- nucleotide repeats such as (AT)₈.
CC (GT)₁₀, (CT)₁₀ and such like as probes. 70 G; 92 T;
SQ Sequence 354 BP; 1.1% Score 14; DB 18; Length 354;

Query Match 1.3% Score 14; DB 20; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 229 cttctctggagccca 242

Qy 181 cttctctggagccca 196

RESULT 112

ID Q55233 standard; DNA: 359 BP.
DE 20-AUG-1993 (first entry)
DE Human small lung cell carcinoma clone SCLC-A.
RW SCLC; calcium channel; ss.
OS Key: espiens. Location/Qualifiers
FH Key misc_feature 12..359
FT misc_feature /*tag= a
FT misc_feature /*tag= b
FN M9401580-A.
PD 20-JAN-1994.
PF 30-JUN-1993; AUG320.
PI (GSR) COMONWELTH SCI & IND RES ORG.
PI Scott NS, Thomas ME;
DR WPI: 94-035083/04.
RT Revi ribosome DNA probe sequences - for the accurate identification
P6 Claim 26: Fig 5c: 55pp; English.
CC The sequences given in Q55231-50 are derived from ribosomes of the
CC grapevine genus vitis. These sequences represent the IGS region of
CC the ribosomal rDNA repeat and contain polymorphisms. These
CC sequences are derived from different grape cultivars. In these clones
CC sequences and were identified in a genomic library of grapevine DNA
CC using simple di-, tri- or tetra- nucleotide repeats such as (AT)₈.
CC (GT)₁₀, (CT)₁₀ and such like as probes. 70 G; 92 T;
SQ Sequence 354 BP; 1.1% Score 14; DB 18; Length 354;

Query Match 1.3% Score 14; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.47e+01; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 50 cttctctggagccca 63

Qy 181 cttctctggagccca 196

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 229 cttctctggagccca 242

Qy 181 cttctctggagccca 196

RESULT 112

ID Q55233 standard; DNA: 360 BP.
DE 08-APR-1996 (first entry)
DE Rat insulin-II gene extended 5'-flanking region.
RW Rat insulin; tissue-specific gene expression; promoter;
RW Human cloning; mapping; non-biased library; diagnosis; detection;
OS Key: espiens. Location/Qualifiers
FH Key misc_feature 12..359
FT misc_feature /*tag= a
FT misc_feature /*tag= b
FN M9401580-A.
PD 20-JAN-1994.
PF 30-JUN-1993; AUG320.
PI (GSR) COMONWELTH SCI & IND RES ORG.
PI Scott NS, Thomas ME;
DR WPI: 94-035083/04.
RT Revi ribosome DNA probe sequences - for the accurate identification
P6 Claim 26: Fig 5c: 55pp; English.
CC The sequences given in Q55231-50 are derived from ribosomes of the
CC grapevine genus vitis. These sequences represent the IGS region of
CC the ribosomal rDNA repeat and contain polymorphisms. These
CC sequences are derived from different grape cultivars. In these clones
CC sequences and were identified in a genomic library of grapevine DNA
CC using simple di-, tri- or tetra- nucleotide repeats such as (AT)₈.
CC (GT)₁₀, (CT)₁₀ and such like as probes. 70 G; 92 T;
SQ Sequence 360 BP; 1.1% Score 14; DB 18; Length 360;

Query Match 1.3% Score 14; DB 17; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Db 36 cttctctggagccca 49

Qy 152 cttctctggagccca 165

[illegible]

Page 122

```

WYTR locus M18917 flanking regions.
DB      Homo sapiens.
OS      Human.
FF      Repeat tandem repeat; Oligonucleotide probe.
FF      Key
FF      Location/Qualifiers
FF      repeat_region
FF      234..404
FF      /note="see comment"
FF      /tag= b
FF      /number= 1
FF      repeat_unit
FF      234..290
FF      /tag= c
FF      /number= 2
FF      repeat_unit
FF      348..404
FF      /tag= d
FF      /number= 3
FF      repeat_unit
FF      405..461
FF      /tag= e
FF      /number= 4
FF      W09110748.A.
PD      23-JUL-1991. D00196.
PD      09-JAN-1991. U00196.
PD      09-MAR-1991. X00196.
PI      (COLA ) COLLABORATIVE REAGS INC.
PI      Keith TP, Mao J, Rose SD;
DR      WPI: 91-238036/32.
DR      Probe and primers from specific human genomic loci - used for
DR      screening DNA samples in forensic applications.
DR      Disclosure: Page 23; 49pp; English.
PS      The sequence was detected by probe CR1-IJ59-2 and comprises the 5'
PS      flanking region and the first 4 tandem repeat units of a section of
PS      chromosome 11p15.5. The last 4 tandem repeats are given in Q1316.
PS      The last 4 tandem repeats are given in Q1316.
CC      Additional unsequenced repeats are present between these two
CC      conditions. See also Q12905-Q12913 and Q13611-Q13617.
CC      Sequence 404 BP; 100 A; 143 C; 71 G; 90 T;
CC      Query Match 1.3%; Score 15; DB 2; Length 404;
CC      Best Local Similarity 100.0%; Pred. No. 8.3%+0.1;
CC      Matches 15; Conservative 0; Mismatches 0; Caps 0;
DB      74 qcacttcgcatggctt 88
DB      |||
DB      |||
CP      327 gcacgtccatgcctt 313

RESULT 118
ID   N80626 standard; DNA; 413 BP.
AC   N80626; 1990 (first entry)
CD   Bovine Rong Ribonuclease Protein-3 in bp-819 detected by probe #2.

```

FT cds 41..373
FT /tag= a
FT W0910409-A.
FT GENE-1 Genentec Institute, Inc.
FT 07-APR-1988; US-179197.
FT (GENE-1) Genentec Institute, Inc.
FT Wang SA, Wozney JM, Rosen V,
FT P-250B; P33213.
FT Isolated DNA sequence and purified BMP-3 protein prep. - by culturing
FT cell transformed with defined DNA and recovery of 96 amino acid protein.
FT (Accession: P0340; 31pp; English).
FT The sequence was isolated from clone Lambda BP-819 by
FT PCR amplification of the DNA sequence.
FT CDS of N92200. The sequence was isolated from clone Lambda BP-819 by
FT hybridization to a probe (N93420). Purified BMP-3 may be used to induce
FT bone and/or cartilage formation and to promote wound healing and tissue
FT repair.
SQ Sequence 413 BP: 119 A; 95 C; 111 G; 88 T;
Query Match 1.34; Score 14; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 138 agagtcagcaaca 151
Cp 269 AAGATGACACGAC 256
RESULT 120 standard; cDNA; 413 BP.
ID T68219
AC N92200
DE 06-APR-1990 (first entry)
DE Toxoplasma gondii p28 gene fragment.
DE P28 antigen; vaccine; diagnosis; toxoplasmosis; hybridization;
DE Probe; antigen; vaccine; diagnosis; toxoplasmosis; hybridization;
DE Toxoplasma gondii RH strain.
PS DS5613135-A.
PD 27-MAY-1997.
PF 03-NOV-1988; 431578
PF 03-NOV-1988; 431578
PR 03-FEB-1992; US-244551
PR 07-JUN-1995; US-477357
PA (REPA-1) RES INST PALO ALTO MEDICAL FOUND.
PI De Araujo, G; Prince JB, Remington JS, Sharma SD;
PT Nucleic acid encoding Toxoplasma gondii P28 protein - for producing
PT recombinant protein for vaccination or diagnosis of infection e.g.
PT in AIDS patients
PS Claim 11: Column 17-18, 10pp; English.
SQ A polynucleotide (T68219) comprises a portion of the coding sequence

CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC cloning and/or mapping of the gene.
SQ Sequence 422 BP: 91 A; 75 C; 101 G; 138 T;
Query Match 1.34; Score 14; DB 21; Length 422;
Best Local Similarity 91.3%; Pred. No. 1.47e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 14 caggtgtcaagag 28
Cp 337 CAGGTGTCACAGG 341
RESULT 122 standard; cDNA; 446 BP.
ID G55114
AC T68219
DE 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST00357.
DE Gene transcription product; genetic markers; tagging; in vivo;
DE Transcription; mapping; locations; chromosomes; chromosomal; se.
PS W09316178-A.
PD 19-AUG-1993.
PF 12-FEB-1993; D01294.
PF 12-FEB-1993; D01294.
PR (US92) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ;
PT WPI; 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes.
PS Example 1; Page 123; 500pp; English.
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for gene identification. The ESTs are derived from a cDNA library
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prep. of antisense sequences, probes and constructs.
CC The ESTs are derived from a cDNA library which has been screened
CC identified as being associated with Klesin. It shows strong
CC sequence similarity to the squid (67%) and Drosophila (70.4%)
CC Klesin heavy chain. See also 059041-061440.
SQ Sequence 446 BP: 82 A; 116 C; 103 G; 144 T;
Query Match 1.34; Score 15; DB 8; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW Bone morphogenetic protein; BMP-3; probes; cartilage formation;
KW bone formation; osteogenic cpds.; prodontal disease; se.
PS Koz taurus.
PS cds
FT Location/Qualifiers
FT 41..374
FT /tag= a
FT /product=BMP-3
FT /tag= b
FT /product-typtic fragment 12
FT 322..352
FT /tag= c
FT /product-typtic fragment 10
PD W0800205-A.
PD 14-JAN-1988.
PF 30-JUN-1987; 001537.
PF (GENE-1) Genentec Institute, Inc.
PF Wozney JM, Rosen V.
PI Wozney JM, Rosen V.
DR P-250B; P33213.
DR Isolated DNA sequence and purified BMP-3 protein prep. and used
DR for inducing cartilage and bone formation.
PS Disclosure; 7pp; English.
CC This region of bp-819 hybridizes to probe 42. The bp-819 clone
CC encodes the bovine bone growth factor BMP-3.
CC and N0627. The peptide is 175 amino acids long. bp-819 of N06026
CC and N0627. The peptide is 175 amino acids long. bp-819 of N06026
CC This sequence is also used to screen a human genomic library to
CC detect the BMP-3 human bone inductive factor which is significantly
CC homologous to N06026 and N06027.
SQ Sequence 413 BP: 119 A; 95 C; 111 G; 88 T;
Query Match 1.34; Score 14; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 138 agagtcagcaaca 151
Cp 269 AAGATGACACGAC 256
RESULT 119 standard; cDNA; 413 BP.
ID N92200
AC N92200
DE 06-APR-1990 (first entry)
DE Toxoplasma gondii p28 gene fragment.
DE P28 antigen; vaccine; diagnosis; toxoplasmosis; hybridization;
DE Probe; antigen; vaccine; diagnosis; toxoplasmosis; hybridization;
DE Toxoplasma gondii RH strain.
PS DS5613135-A.
PD 27-MAY-1997.
PF 03-NOV-1988; 431578
PF 03-NOV-1988; 431578
PR 03-FEB-1992; US-244551
PR 07-JUN-1995; US-477357
PA (REPA-1) RES INST PALO ALTO MEDICAL FOUND.
PI De Araujo, G; Prince JB, Remington JS, Sharma SD;
PT Nucleic acid encoding Toxoplasma gondii P28 protein - for producing
PT recombinant protein for vaccination or diagnosis of infection e.g.
PT in AIDS patients
PS Claim 11: Column 17-18, 10pp; English.
SQ A polynucleotide (T68219) comprises a portion of the coding sequence

CC (see also T68219) for the p28 antigen (M1815) of Toxoplasma gondii.
CC The coding sequence was deduced from cDNA clones isolated from a
CC cDNA library by screening with mouse anti-P303-Ag as probe and
CC clones library by hybridization using isolated P303-Ag cDNA clones as
CC probes. The coding sequence was deduced from cDNA clones isolated from
CC at least 10 consecutive nucleotides of the sequence, can be labelled
CC with a detectable tag and used in hybridization assays for the
CC detection of Toxoplasma infections e.g. in AIDS patients.
SQ Sequence 413 BP: 103 A; 95 C; 105 G; 106 T;
Query Match 1.34; Score 14; DB 30; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 397 aaagtcagagtg 410
Cp 193 AAAAAACGACGTC 180
RESULT 121 standard; cDNA to mRNA; 422 BP.
ID T24490
AC T24490
DE 22-SEP-1996 (first entry)
DE Human gene signature; messenger RNA; mRNA; relative abundance; frequency;
DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;
DE human; cloning; mapping; non-biased library; diagnosis; detection;
DE cell typing; abnormal cell function; se.
PS W09316178-A.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PF 12-NOV-1993; J01916.
PR (US94) US DEPT HEALTH & HUMAN SERVICE.
PI Matubara K, Ohubo K;
PT WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for gene identification. The ESTs are derived from a cDNA library
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues.
PS Claim 1; Page 1622; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC cDNA) was used to identify gene signatures in a 3'-directed cDNA
CC library. The cDNA library was constructed from a cDNA library
CC given in T3001-T24837 and which is able to hybridize to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3' end of the cDNA. The cDNA library was constructed from a cDNA
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of

Db 273 gctctctctctac 287
QY 789 GCTCTCTCTCTAC 803

RESULT 123
ID Q39702 standard: DNA: 446 BP.
AC Q39702:
DE 20-MAY-1993 (first entry)
KW Expressed sequence tag human gene marker EST00257.
KW human gene sequencing; PCR mapping; somatic cell hybrids;
KW sublocalization; gene tagging; tissue typing.
OS Synthetic.
PS 07-JUN-1993 A.
PF 19-JUN-1993: 005222.
PR 20-JUN-1993: 05-716831.
PR 12-FEB-1995: 05-837195.
PA LQSHB US DEPT HEALTH & HUMAN SERVICE.
WP: 93-036325/04 C.
DR Particular expressed sequence tags from human CDNA - corresponds
PT to transcription prod. of genes, useful for tagging genes.
PS Cloning; tissue typing
PS Cloning; tissue typing
CC This sequence represents an EST (expressed sequence tag) ESTs are markers
CC for human genes actually transcribed in vivo. Unlike the random genomic
CC DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.
CC They are useful for identifying and cloning genes of interest, and for
CC sequencing within a few years at fractionation of most expressed human
CC genes. Using PCR primers Q39419-Q39560 (sequences designed
CC from the ESTs) sublocalization of an EST can be achieved with panels of
CC fragments from specific chromosomes or pools of large genomic clones in
CC a Southern blot. This sequence represents EST00257.
SQ Sequence 446 BP; 82 A; 116 C; 103 G; 144 T.

Query Match 1.38; Score 15; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Caps 0;
Matches 15; Conservative 0; Mismatches 0;

Db 273 gctctctctctac 287
QY 789 GCTCTCTCTCTAC 803

RESULT 124
ID T11846 standard: DNA: 468 BP.
DE 03-SEP-1996 (first entry)

DE P. putida LMG 2232 16S-23S rRNA spacer region.
KW Probe; detection; identification; microorganisms; amplify;
KW 16S-23S rRNA spacer region; respiratory tract; universal;
KW Pseudomonas putida.
OS Pseudomonas putida.
PS WO9600298-A1.
PD 04-JAN-1996: E02452.
PR 24-JUN-1994: EP-870106.
PR 07-APR-1995: EP-870032.
PA (INNO-) INNOGENETICS NV.
DI Janne C, Rosau X, Van Heuverswyn H.
PT Novel hybridisation assay for the detection of eubacteria - esp
PT Mycobacterium, comprises amplifying the 16S-23S rRNA spacer region
CC Claim 2: Fig 40; 24pp: English.
CC The sequences given in 111791-857 represent the 16S-23S rRNA spacer
CC region of Mycobacterium species. The method comprises amplifying
CC at least one or more microorganisms. The method comprises amplifying
CC the 16S-23S rRNA spacer region and hybridising one or more of the probes
CC given in 111791-857 to the amplified sequence. These probes were
CC derived from the sequences of the spacer region of Mycobacterium
CC putida LMG 2232. This spacer region is derived from Pseudomonas
SQ Sequence 468 BP; 117 A; 101 C; 117 G; 133 T.

Query Match 1.38; Score 14; DB 21; Length 468;
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
Matches 14; Conservative 0; Mismatches 0;

Db 338 tatctctctctac 351
QY 1092 TATCTCTCTAC 1105

RESULT 135
ID T59835 standard: DNA: 494 BP.
AC T59835:
DE 22-APR-1997 (first entry)
KW Beta-casein gene fragment #7, antithrombin III; factor XI; factor X;
KW factor XIII; mammary gland specific; hereditary AT3 deficiency; therapy.
KW acquired AT3 deficiency; heparin affinity; as.
OS Capra hircus.
PS Location/Qualifiers
FT misc_feature 1
FT /tag= a
FT /note= unspecified number of bases, joins to n at
FT position 156 of T59834.

intron 1..72
FT /tag= b
FT /number= 8
FT /note= fragment of intron 8
FT /tag= c
FT /number= 9

WO96268-A1.
DE 23-DEC-1996: U03420.
PR 21-FEB-1995: US-391743.
PA (GENE) GENETICS TRANSGENICS CORP.
DI Cole ES, Dittulio P, Meade H;
PT New transgenic antithrombin III - useful for treating
PT acquired or inherited AT3 deficiency, with faster clearance rate
PT than plasma AT3
PS Diclosure; Fig 10; 37pp: English.
CC between them encoded the protein represented by WO96268-A1.
CC used to produce the transgenic antithrombin III (AT3) of the invention.
CC AT3 is a serine protease inhibitor, which inhibits thrombin and the
CC activated forms of factors X, VII, IX, XI, and XII. The transgenic AT3
CC N-acylgalactosamine (GALNAc) monomer. To produce a mammary gland
CC transgene, human AT3 cDNA (cloned as an 18.5 kb fragment) was inserted
CC into the goat beta-casein gene. The final 14.95 kb vector was
CC injected into goat embryos, which when developed produced the AT3
CC glycoprotein. The glycoprotein was purified from the milk of the
CC goats. The glycoprotein has a complex oligosaccharide at each glycosylation site
CC except the Asn155 residue. At Asn155, it has a significant amount of
CC oligomannose, and hybrid residues. The AT3 is used in the same way as
CC plasma AT3 for the treatment of hereditary or acquired AT3 deficiency.
CC also have a higher affinity for heparin. The AT3 can also be produced
CC without the variability, immunogenicity and viral contamination problems
CC associated with plasma-derived material.
SQ Sequence 494 BP; 176 A; 82 C; 62 G; 193 T.

Query Match 1.38; Score 15; DB 28; Length 494;
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Caps 0;
Matches 15; Conservative 0; Mismatches 0;

Db 282 aaatcatgacatc 296
CP 669 AAATCATGACATC 555

RESULT 126
ID Q14952 standard: DNA: 539 BP.
AC Q14952:

DE 17-JUN-1992 (first entry)
KW Bovine epithelium precursor? (partial).
KW ET; growth regulation; inhibition; stimulation; as.
OS Bos taurus.
PS WO915510-A.
PD 03-APR-1991: 002121.
PR 03-APR-1990: US-504508.
PR 13-MAR-1991: US-083796.
PA (BIRM) BRISTOL-MYERS SQUIBB.
DI Smith, William G.
WP: 91-325168/44.
DR P-PS08; R15426.
PT New cysteine-rich growth modulating proteins, epithelins - useful
PT as inhibitors of neoplastic cell growth and to promote wound
PT healing and tissue repair.
PS Diclosure; Fig 24; 97pp: English.
CC The epithelins appear to comprise several distinct members sharing
CC significant structural homology. Two members of the epithelin family,
CC ET-1 and ET-2, have been purified from human endothelial cells and only
CC encoding these and seven other members of the epithelin family have
CC been isolated from rat (Q14339), human (Q14339), bovine, murine
CC (Q14340) and chicken (Q14853).
CC ET-1 is a bifunctional growth regulator, capable of stimulating
CC ET-2 is functionally similar to ET-1, but inhibits growth of others.
CC ET-2 is functionally similar to ET-1, but inhibits growth of others.
CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonises this ET-1 activity.
CC Q14339, Q14340, Q14853, and R15315-20.
SQ Sequence 539 BP; 99 A; 163 G; 106 T.

Query Match 1.38; Score 14; DB 3; Length 539;
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
Matches 14; Conservative 0; Mismatches 0;

Db 48 cctctctctctctc 61
QY 891 CCTCTCTCTCTCTC 904

RESULT 127
ID Q5054 standard: DNA: 546 BP.
DE 18-JUN-1994 (first entry)
KW Sequence of orphan receptor DNA fragment PCR 9
KW PCR: oligo: primer: peptide ligand receptor: as.
OS Synthetic.
PS Location/Qualifiers
FT misc_feature 1..26
FT /tag= a

CC determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
CC also offer protection against a wider range of pathogens.
SQ Sequence 566 BP; 138 A; 114 G; 176 T;
Query Match 1.38; Score 14; DB 19; Length 566;
Best Local Similarity 91.34; Pred. No. 3.47e+01; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 460 cacttttggtaag 474
|||||
Cp 539 TATTTCCTGTTGAG 525

RESULT 136
ID T3132 standard; DNA: 590 BP.
AC T3132 standard; DNA: 590 BP.
DE Exon 7 of gene of IMAGE clone 264611.
KW IMAGE clone 264611; gene fragment; human; chromosome 10; D10S541 marker;
D10S215 marker; tumor suppressor gene; prostatic cancer; cancer therapy;
KW hyperplasia; 10q loss tumor; as.
OS Homo sapiens.
PN 03-JAN-1994; US-181271.
PP 03-JAN-1994; US-181271.
PR 03-JAN-1994; US-005840.
PA (MCR) IMPERIAL CANCER RES TECHNOLOGY.
DR WPI: 97-259037/33.
PT Nucleic acid hybridizing to chromosome 10 tumor suppressor gene -
PT useful for diagnosis, prognosis and treatment of prostatic cancer
PT and for assessing susceptibility to cancer
PT Nucleic acid hybridizing to chromosome 10 tumor suppressor gene -
CC (see T3132 for gene cDNA sequence). These sequences are all nucleic acid
sequences of the invention. The nucleic acid of the invention (I) is able
to hybridize to a region of human chromosome 10 bounded by
the D10S541 and D10S215 markers. The nucleic acid of the invention (I) is
particularly involved in prostatic cancer but also in melanoma, glioma
and non-Hodgkin's lymphoma. Any nucleic acid that hybridizes selectively
to the specified chromosomal region can be used to determine
the presence of prostatic cancer, cancer and/or lymphoma/prostate,
especially of prostatic cancer, cancer and/or lymphoma/prostate,
CC wild-type suppressor gene can also be used to treat cancer, especially
CC when included in a viral vector. Similar detection methods can be based
CC on the amount of protein encoded by (I), or its truncation or loss, in a
CC cell or tissue sample. The protein encoded by (I) is a tumor suppressor
CC the protein, particularly antibodies. The labeled molecule when coupled

CC to a cytotoxin can be used for cancer treatment. The encoded protein can
be used to raise antibodies and these used to screen DNA expression
libraries or for polypeptide isolation. (I) allows differential diagnosis
CC in the detection of prostatic cancer, cancer and/or lymphoma/prostate,
CC have lost this region) and determination of micro-metastases in the
CC blood.
SQ Sequence 590 BP; 201 A; 70 C; 104 G; 212 T;
Query Match 1.38; Score 14; DB 36; Length 590;
Best Local Similarity 100.04; Pred. No. 3.47e+02; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 22 tttgttttttttttt 35
|||||
Cp 542 TTTGATTTTTCGTT 529

RESULT 137
ID Q99783 standard; DNA: 653 BP.
AC Q99783 standard; DNA: 653 BP.
DE 24-JUN-1996 (first entry)
KW SAR: tobacco; protein-synthesis independent gene; cyclohexamide;
KW systemic acquired resistance response; pl.4.3; anti-pathogen;
KW plant protection; as.
OS Nicotiana glauca.
PN 20-JUL-1995.
PP 03-JAN-1994; US-181271.
PR 03-JAN-1994; US-181271.
PA (CIMA) CIMA GENET AG.
DR WPI: 95-26387/24.
PT New DNA contg. plant systemic acquired resistance genes - and
PT transgenic plants contg. them. impart disease and pest resistance,
PT Nucleic acid hybridizing to tobacco genome.
PT Nucleic acid hybridizing to tobacco genome.
CC This sequence represents the DNA sequence of a tobacco protein-synthesis
independent gene. The gene is involved in the regulation of the systemic
acquired resistance (SAR) response. This gene is designated pl.4.3. This
gene encodes a protein of 653 amino acids. The protein is a member of the
recombinant/chimeric DNA molecules of the invention. Q99783-Q99790 and
Q99806 are SAR genes used in the recombinant/chimeric DNA molecules of
the invention. The wild type genes corresponding to these sequences can
all be chemically induced in a plant in a protein-synthesis independent
manner. The genes are involved in the regulation of the systemic acquired
resistance response. These sequences were isolated by differential screening of
a cDNA library, followed by analysis by Northern hybridisation to RNA in
the presence and absence of cyclohexamide. The genes are used in the
creation of transgenic plants. Transgenic expression of 2 or more of the
recombinant molecules of the invention that encode anti-pathogenic

CC proteins provides a synergistic increase in plant protection, and may
CC also offer protection against a wider range of pathogens.
SQ Sequence 653 BP; 165 A; 114 G; 157 G; 215 T;
Query Match 1.38; Score 15; DB 18; Length 653;
Best Local Similarity 100.04; Pred. No. 8.37e+01; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 314 caactttttttttt 328
|||||
Cp 515 CAACTTTTCCTGCA 529

RESULT 138
ID T36496 standard; cDNA: 666 BP.
AC T36496 standard; cDNA: 666 BP.
DE 24-JUN-1997 (first entry)
KW GST: glutathione-S-transferase IIIC-encoding cDNA.
KW GST: glutathione-S-transferase; maize; gene therapy; transgenic;
KW plants; increased herbicide resistance; hetero-oxycetamides; as.
OS Zea mays.
PN 23-JAN-1995; DE-001840.
PP 23-JAN-1995; DE-001840.
PR (FABE) BAYER AG.
PA Steeler B, Hain R, Mann K, Reif H, Reinher P;
PI Thomak JE;
DR P-008; 98820.
PT Increase resistance of plants to hetero-oxycetamide - used to
PT increase resistance of plants to hetero-oxycetamide - used to
PT herbicides.
CC This sequence is a cDNA sequence derived from maize that encodes glutathione-
S-transferase IIIC (GSTIIIC). The gene imparts a natural resistance
in plants to herbicides, esp. of the hetero-oxycetamide type, through
production of the GSTIIIC detoxifying enzyme in vivo. The gene was
isolated from a cDNA library and is a member of the glutathione-S-
transferase gene family. The vector may be used to transform Agrobacterium
tumefaciens, which in turn may be used to infect almost any plant, so
that the GSTIIIC gene is expressed in the plant. Suitable plants that may
be transformed include potato, sugar cane, sugar beet, cereals and
SQ Sequence 666 BP; 103 A; 263 G; 201 G; 99 T;
Query Match 1.38; Score 15; DB 26; Length 666;
Best Local Similarity 100.04; Pred. No. 8.37e+01; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 87 ccagactctctg 100
|||||
Cp 886 CCAGACTCTCTG 873

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 215 accgttactctgca 229
|||||
Cp 398 ACCGTCATCTGCTCA 412

RESULT 139
ID Q68331 standard; DNA: 672 BP.
AC Q68331 standard; DNA: 672 BP.
DE 17-FEB-1995 (first entry)
KW Wheat germin coding sequence.
KW oxalic acid degradation; oxalate oxidase; selection marker;
KW Triticum aestivum.
OS Triticum aestivum.
PN 17-FEB-1995.
PP 17-FEB-1995.
PR (SNPT) ELF SANPT.
PA (SNPT) SOC NAT ELF AQUITALINE.
PI Gress-Besset B, Gresson R, Pignard A, Schneider M;
DR WPI: 94-21680/26.
PT Use of gene encoding oxalic acid degrading enzyme - for plant
cell selection, esp. gene coupled to gene-encoding protein
conferring pathogen resistance
CC This sequence is a cDNA sequence derived from wheat, a protein induced during
germination of wheat, is a preferred example of DNA coding for a
protein able to degrade oxalic acid. Such DNA can be coupled to a
sequence of interest, e.g. one coding for enzymes which confer
resistance to pathogens (esp. endochitinases). Plant cells which
express the coupled sequence are selected and grown on a medium
containing oxalic acid. The wheat
germin sequence was published in J.Biol.Chem. 264, 4896-4900 (1989)
and in J.Biol.Chem. 268, 10461-10469 (1991).
SQ Sequence 672 BP; 122 A; 235 C; 188 G; 127 T;
Query Match 1.38; Score 14; DB 11; Length 672;
Best Local Similarity 100.04; Pred. No. 3.47e+02; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 87 ccagactctctg 100
|||||
Cp 886 CCAGACTCTCTG 873

[illegible]

```
Query Match      1.38; Score 14; DB 32; Length 688;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 141
 ID T79657 standard; DNA: 689 BP.
 AC T79657:
 DT 04-FEB-1998 (first entry)
 DE BRCA2 cancer susceptibility gene exon 29a.
 KW BRCA2 cancer susceptibility gene; breast cancer;
 KW breast cancer; colorectal cancer; ovarian cancer;
 KW leukodystrophy; prostate cancer; retinoblastoma;
 OS Homo sapiens.

Tue Nov 17 08:55:26 1998

US-08-887-977-9.IND

[illegible]

```

Query Match      1.3%   score 14;  DB 22;  Length 702;
Best Local Similarity 100.0%;  pred. No. 3.47e+02;
Matches 14;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db      230 agcttctctatcaca 243
      |||||
Cp      1107 AGGCTTCTCTACACA 1094

```

	RESULT	I43	
	ID:	T35193	standard; DMA: 713 BP.
		T35199;	AC
	DT	28-Nov-1996	(first entry)
	DE	Mouse interleukin-12 p35 subunit gene.	
	DEFN	P40 subunit fusion protein; Interleukin-12, IL-12; p35 subunit;	
	EVAL	p40 subunit fusion protein; Interleukin-12, IL-12; p35 subunit;	
	KW	fibrosarcoma; renal cell carcinoma; immunotherapy; therapy;	
	REV	retrovirus; viral ds.	
	OS	Mus musculus.	
	Key words	Location/Qualifiers	
	Cds	/tag= a	
	Ff	signal_peptide	47..112
	Ft	/tag= b	
	mat_peptide		113..591

GB23074077-9A.
28-MAY-1996; 024453.
29-NOV-1996; 024453.
29-NOV-1996; GB-021961.
29-NOV-1996; GB-021961.
14-DEC-1995; GB-025555.
(CANC.) CANCER RES CAMPAIGN TECHNOLOGY.
(TUDU.) UNIV DUKE.
Adsworth, Ailsa, Pittal PA, Stratton MR, Wooster RP.
Nucleic acid molecules comprising part or all of the BRCA2 cancer
susceptibility gene useful for diagnosis, prognosis or therapeutic
treatment of cancer.
Claim 1, Fig 2, 124pp; English.
The BRCA2 gene, the genomic organization of the BRCA2
cancer susceptibility genes exon 20a. The nucleic acid molecules can be
used to construct probes for screening cDNA or genomic libraries,
sequencing positive clones obtained, and assembling the full length
BRCA2 sequence. The BRCA2 nucleic acid molecules and proteins are
usefully for medical treatment. Preferably gene therapy,
breast cancer, ovarian prostate or colorectal cancer, ocular melanoma
or Leukemia. In particular antisense oligonucleotides capable of
hybridizing to the BRCA2 nucleic acid, pre-mRNA or mature mRNA are used
so that the expression of the BRCA2 nucleic acid is reduced or
abolished. The BRCA2 nucleic acid molecules are useful for
diagnosing susceptibility or predisposition to cancer in a patient.
Nucleic acid molecules are used to design probes or primers for PCR to
determine or detect the presence of mutations in a sample of nucleic
acid from a patient. The BRCA2 promoter region is useful for screening
for mutations in the promoter region. The BRCA2 nucleic acid molecules
control of the promoter. Antibodies are used to determine the presence,
amount or location in a cell of a BRCA2 polypeptide or its mutant forms.
The polypeptides are used to screen for binding partners, these are
useful to screen for substances which mimic the activity of BRCA2
protein. 689, 897, 204, 132, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 8

```

Query Match      13%  Score 14:  DB 35:  Length 689;
Best Local Similarity 100.0%:  Pred. No. 3.47e+02;
Matches 14:  Conservative 0:  Mismatches 0:  Indels 0:  Caps 0:

DBD      257 tgaagattattttc 270
          |||||
          17 tgaagattattttc 30

RESULT 142
ID      T32146 standard; cDNA; 702 BP.
DEF     T32146 standard; cDNA; 702 BP.
DT      07-NOV-1996 (first entry)

```

Page 156

119-08-887-977-9

1509242475-AL.
 W-PM-1996.
 07-PM-1996; 001787;
 08-FEB-1995; US-385335.
 (WHOED) WHITEHEAD INST BIOMEDICAL RES.
 Lierchke GJ, Mulligan SC;
 P-PM-1996-398.
 P-PM-199246.
 New DNA encoding fusion protein. exp. conty. IL-12 p15 and p40
 subunits - for treatment of established tumours or prevention of
 tumour establishment
 Example 1, Page 33-54; 118pp: English
 subunit (R89246) of murine Interleukin-12 (IL-12). It can
 be utilised in novel constructs in which the p15 coding sequence
 is joined to the IL-12 p40 coding sequence (see also T35200) via a
 DNA fragment (T35193-97 and T35201-03) coding for a peptide linker
 vector (see also T35198) allowing construction of bioactive dimeric
 IL-12 in transfected cells. CHS-5, B16 or renal carcinoma
 cells secreting the IL-12 dimer can be used to reduce the size of
 established tumours and/or increase survival time, exp. in cases
 of melanoma, 711, fibrosarcoma and renal cell carcinoma. 151. 7

```
Query Match      1.38; Score 14; DB 23; Length 713;  
Best Local Similarity 100.0%; Pred. No. 3,47+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

RESULT 146
# 751 BP standard; DNA, 751 BP.
Q8B156 LC
Q8B156 IC
20-FEB-1995 (first entry)
Wheat grain genomic clone gf-2.8.
What germ degradation: oxalate oxidase; selection marker: germin;
oxalate oxidase activity: 100%; oxalate tolerance: 100%;
Oxidation resistance: 100%; hydrogen resistance; plant cell selection; ab.
Triticum aestivum Location/Qualifiers
Key cda
      /tag= "a"
      /note= "germin"
      /tag= "c"
      /note= "clone gf-2.8"
WG94L3790-A.
23-JUN-1994. F01003
07-DEC-1993. F01003
07-DEC-1993. F01003

```

PA (SNT) ELP SANOFI.
 PI Genter, Beest B, Gilson R, Pignard A, Schneider M;
 DR WP: 94-217870/26.
 CC Use of gene encoding oralic acid degrading enzyme - for plant
 CC conferring pathogen resistance.
 CC The sequence coding for Wheat germ, a protein induced during
 CC germination of wheat. Is a preferred example of DNA coding for a
 CC resistance to pathogens (esp. endochitinases). Plant calli which
 CC have been successfully transformed by the sequences are selected
 CC germ sequence was published in J Biol. Chem. 228, 10461-10469
 CC (1991).
 CC Sequence 751 BP: 149 A; 249 C; 200 G; 153 T;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 CC Db 107 cagagactctctg 120
 CC Cp 886 cagagactctctg 873
 CC
 CC RESULT 145
 CC ID T94511 standard; DNA: 756 BP.
 CC AC T94511: (first entry)
 CC DE Toxoplasma gondii p28 coding sequence.
 CC KW infection; vaccine; immunosay; antibody; toxoplasmosis; ss.
 CC OS Toxoplasma gondii.
 CC PR Key Location/Qualifiers
 CC FT CDS 1..756
 CC FT /product= "p28 protein"
 CC FT /note= "no STOP codon given at end of sequence"
 CC FT US6565542-A.
 CC PR 03-FEB-1998; 264951.
 CC PR 03-NOV-1989; US-431578.
 CC PA (REPA.) RES INST PALO ALTO MEDICAL FOUND.
 CC PI De Araujo FG, Prince JB, Remington JS, Sharma SD;
 CC DR WP: 97-456789/42.
 CC DR P-PDB: N36012.
 CC

PT Nucleic acid encoding Toxoplasma gondii p28 protein - used in
 PT hybridization assays for diagnosing toxoplasmosis, also for
 PT expressing proteins in vaccines and diagnostic reagents
 PT Claim 14; Column 19-20; 14pp; English.
 CC This nucleotide sequence corresponds to the coding sequence encoding
 CC the Toxoplasma gondii p28 protein. The sequence was obtained from a
 CC library generated from nucleic acid isolated from T. gondii tachyzoites.
 CC CC i.e. toxoplasmosis. The encoded protein is useful in vaccines and
 CC probes in hybridization tests for direct detection of T. gondii infection
 CC i.e. toxoplasmosis. The encoded protein is useful in vaccines and
 CC diagnostic (immunosay) compositions for detection of anti-T. gondii
 CC antibodies. Antigenic peptides from the p28 protein can be used to raise
 CC antibodies. 756 BP: 201 A; 192 C; 208 G; 155 T;
 CC Query Match 1.3%; Score 14; DB 35; Length 756;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 CC Db 506 aaaaagcaagctg 519
 CC Cp 193 aaaaagcaagctg 180
 CC
 CC RESULT 146
 CC ID T94511 standard; DNA: 756 BP.
 CC AC T94511: (first entry)
 CC DE Toxoplasma gondii p28 gene coding sequence.
 CC KW p28 antigen; vaccine; diagnosis; toxoplasmosis; ss.
 CC OS Toxoplasma gondii.
 CC PR Key Location/Qualifiers
 CC FT mRNA 110..522
 CC FT /tags= "a"
 CC FT /note= "Claim 11"
 CC FT US633139-A.
 CC PR 27-MAY-1987.
 CC PR 03-NOV-1989; 431578.
 CC PR 03-NOV-1989; US-431578.
 CC PR 03-NOV-1989; US-431578.
 CC PR 07-JUN-1995; US-477157.
 CC PA (REPA.) RES INST PALO ALTO MEDICAL FOUND.
 CC PI De Araujo FG, Prince JB, Remington JS, Sharma SD;
 CC DR WP: 97-297357/27.
 CC DR P-PDB: N36012.
 CC PT Nucleic acid encoding Toxoplasma gondii p28 protein - for producing
 CC recombinant protein for vaccination or diagnosis of infection e.g.
 CC in AIDS patients
 CC Claim 2; Column 15-16; 10pp; English.
 CC Claim 2; Column 15-16; 10pp; English.
 CC CC coding sequence for the p28
 CC antigen (W18315) of Toxoplasma gondii. The sequence was deduced

CC from cDNA clones isolated from a cDNA library by screening with
 CC mouse anti-P30-AG as probe and clones obtd. by hybridization using
 CC anti-P30-AG as probe. One isolated clone (see
 CC US 633139) including the p28 coding sequence (see
 CC US 633139) can be used for the prodn. of recombinant p28 polypeptides for
 CC use in diagnosis or vaccines, and/or can itself be directly used in
 CC hybridization assays (see also T94511-21) for the detection of
 CC Toxoplasma gondii in AIDS patients.
 CC Sequence 756 BP: 201 A; 192 C; 208 G; 155 T;
 CC Query Match 1.3%; Score 14; DB 30; Length 756;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 CC Db 506 aaaaagcaagctg 519
 CC Cp 193 aaaaagcaagctg 180
 CC
 CC RESULT 147
 CC ID T94511 standard; DNA: 756 BP.
 CC AC T94511: (first entry)
 CC DE Toxoplasma gondii p28 coding sequence #2.
 CC KW infection; vaccine; immunosay; antibody; toxoplasmosis; ss.
 CC OS Toxoplasma gondii.
 CC PR Key Location/Qualifiers
 CC FT CDS 1..756
 CC FT /tags= "a"
 CC FT /product= "p28 protein"
 CC FT /note= "no STOP codon given at end of sequence"
 CC FT US6565542-A.
 CC PR 03-FEB-1998; 264951.
 CC PR 03-NOV-1989; US-431578.
 CC PR 03-FEB-1992; US-264951.
 CC PA (REPA.) RES INST PALO ALTO MEDICAL FOUND.
 CC PI De Araujo FG, Prince JB, Remington JS, Sharma SD;
 CC DR WP: 97-456789/42.
 CC DR P-PDB: N36012.
 CC PT Nucleic acid encoding Toxoplasma gondii p28 protein - used in
 PT hybridization assays for diagnosing toxoplasmosis, also for
 PT expressing proteins useful in vaccines and diagnostic reagents
 PT Claim 14; Column 19-20; 14pp; English.
 CC This nucleotide sequence corresponds to the coding sequence encoding
 CC the Toxoplasma gondii p28 protein. The sequence was obtained from a
 CC library generated from nucleic acid isolated from T. gondii tachyzoites.
 CC CC fragments of the genomic sequence (T94510) e.g. T94511-5 can be used as
 CC probes in hybridization tests for direct detection of T. gondii infection

CC i.e. toxoplasmosis. The encoded protein is useful in vaccines and
 CC diagnostic (immunosay) compositions for detection of anti-T. gondii
 CC antibodies. Antigenic peptides from the p28 protein can be used to raise
 CC antibodies. 756 BP: 201 A; 192 C; 207 G; 155 T;
 CC Query Match 1.3%; Score 14; DB 35; Length 756;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Caps 0;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 CC Db 506 aaaaagcaagctg 519
 CC Cp 193 aaaaagcaagctg 180
 CC
 CC RESULT 148
 CC ID T10931 standard; DNA: 782 BP.
 CC AC T10931: (first entry)
 CC DE Borrelia burgdorferi exported plasmid protein A (BppA).
 CC KW BppA; Lyme disease; antigenic; immune response; detection; virulent;
 CC infection; ss.
 CC OS Borrelia burgdorferi.
 CC PR Key Location/Qualifiers
 CC FT misc_feature 107..135
 CC FT /tags= "a"
 CC FT /note= "dyad symmetrical sequence, suggests that
 CC the BppA gene is transcriptionally
 CC regulated"
 CC FT -35_signal 114..119
 CC FT /tags= "b"
 CC FT -10_signal 138..143
 CC FT rbs 151..156
 CC FT /tags= "d"
 CC FT /note= "putative ribosome binding site"
 CC FT cde 157..168
 CC FT /tags= "e"
 CC FT signal_peptide 159..218
 CC FT /tags= "f"
 CC FT mat_peptide 219..665
 CC FT primer_bind 219..237
 CC FT /tags= "h"
 CC FT /note= "binding site of primer CC01"
 CC FT primer_bind 646..668
 CC FT /note= "binding site of primer CC03"
 CC W0953114-A1.
 CC PR 28-DEC-1995.
 CC PR 18-JUN-1995; 007740.

17-JUN-1994; US-261825.
PA (BREC) UNIV CALIFORNIA.
PI Bianco DR, Champion CI, Haake DA, Lovett WA, Miller JN;
DR WPI: 96-0528/06.
DR P-P2D; 885801.
CC The human interferon gamma gene has been cloned and sequenced.
PT diagnostic markers for Lyme disease to B burgdorferi and as
PS Claim 6; Page 44-45; Fig 2; 67pp; English.
CC TIGR1 is a genomic DNA sequence from Borrelia burgdorferi, it
CC encodes a protein of about 17 kD. B burgdorferi is a virulent
CC pathogen of humans and animals. The protein encoded by TIGR1
CC induces an immune response to pathogenic Borrelia burgdorferi (BB)
CC in an animal. The sequence may be used to produce recombinant BpA
CC expressed in E. coli. BpA is exported beyond the inner membrane and
CC secreted into the culture medium. The protein encoded by TIGR1
CC growth polynucleotide sequence. BpA and BpA antibodies can
CC be used for the detection of pathogenic BB in a sample.
SQ Sequence 782 BP; 327 A; 77 C; 117 G; 281 T;
Query Match 1.38; Score 14; DB 19; Length 782;
Best Local Similarity 100.00; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 439 tataaagcaaaa 452
Cp 197 TATAAAGCAAAA 184

RESULT 149
AC 003870; standard; DNA: 790 BP.
AT 31-AUG-1990 (first entry)
DE H. influenzae DNA fragment containing the PROMP-2 gene
KW outer membrane proteins; PROMP-2; active immunisation;
OS Homo sapiens. Location/Qualifiers
FH Key
FI Location/Qualifiers
FT 161..628
FT /tag= a
FT /product= PROMP-2, outer membrane protein of H. influenzae.
FN M9090257-A.
PD 22-MAR-1990.
PF 31-AUG-1989; 003179.
PR 31-AUG-1989; US-246572.
PR 21-DEC-1989; US-596274.
PA (PRAX).
PI Anilomlis A, Seid RC, Deich RA, Zlotnick GW, Green BA;
DR WPI: 90-115815/15.
CC Outer membrane protein epitopes of Haemophilus influenzae- used in

PT the prodn. of antibodies, in vaccines and for prodn. of reagents for
PT diagnosis.
PS Disclosure; Fig 14; 16pp; English.
CC The sequence is a 789bp BstEII-XbaI fragment of pA130. The predicted
CC coding region is given above.
CC See also 003869, R03948 and R03949.
SQ Sequence 790 BP; 240 A; 106 C; 176 G; 286 T; 2 Others;
Query Match 1.38; Score 14; DB 1; Length 790;
Best Local Similarity 100.00; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 63 tctgtctgtgttt 76
Cp 749 TCTGCTGTGTTT 762

RESULT 150
ID T94515 standard; DNA: 797 BP.
AC T94515;
AT 12-FEB-1998 (first entry)
DE Toxoplasma gondii p28 genomic sequence.
KW Toxoplasma gondii; p28 tachyzoite probe; hybridisation; detection;
KW Infection; vaccine; immunosay; antibody; toxoplasmosis; as.
OS Toxoplasma gondii.
PR US5685542.2 A.
PR 03-SEP-1989;
PR 03-FEB-1992; 264951.
PR 03-NOV-1989; US-431578.
PR 03-FEB-1992; US-264951.
PA (SEPA-) RES INST PALO ALTO MEDICAL FOUND.
DR WPI: 97-456799/42.
CC Nucleic acid encoding Toxoplasma gondii p28 protein - used in
CC hybridisation assays for diagnosing toxoplasmosis; also for
CC expressing proteins useful in vaccines and diagnostic reagents
CC This nucleotide sequence corresponds to a fragment comprising nucleotides
CC 375-1171 of the genomic sequence (T94510) encoding the Toxoplasma gondii
CC p28 protein. The complete sequence was obtained from a library
CC generated from nucleic acid isolated from T. gondii tachyzoites.
CC Hybridisation tests for direct detection of T. gondii infection
CC i.e. toxoplasmosis. The encoded protein is useful in vaccines and
CC diagnostic (immunosay) compositions for detection of anti-T. gondii
CC antibodies. Antigenic peptides from the p28 protein can be used to raise
CC antibodies.
SQ Sequence 797 BP; 204 A; 188 C; 234 G; 171 T;
Query Match 1.38; Score 14; DB 35; Length 797;
Best Local Similarity 100.00; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 547 aaaaagaagagtg 560
Cp 193 AAAAAGCAAGGCTG 180

RESULT 151
AC 000257; standard; DNA: 810 BP.
AT 28-DEC-1997.
PF 22-OCT-1992; 966121.18.
PR 22-OCT-1992; US-546121.
PA (BOF) BOEHRINGER MANNHEIM GMBH.
PI Baylink DJ, Mohan S;
DR WPI: 98-031813/03.
CC Inhibitory insulin like growth factor binding protein - useful to
PT inhibit effects of insulin like growth factor on bone cells and in
PT diagnostic assays
PS Disclosure; Fig 1; 1pp; English.
CC The sequence encodes the novel inhibitory insulin like
CC growth factor binding protein (In-IGF-BP), which is 25 amino acid
CC residues in length. The In-IGF-BP can be used to modulate the effects
CC of insulin like growth factor I or II (IGF) on bone cells. It is used
CC to treat IGF dependent bone tumors, and in a competitive binding assay
CC to great IGF dependent bone tumors, and in a competitive binding assay
SQ Sequence 810 BP; 135 A; 279 C; 282 G; 114 T;
Query Match 1.38; Score 14; DB 40; Length 810;
Best Local Similarity 100.00; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 541 tctctgcagca 554
Cp 832 TCTCTGCAGCA 845

RESULT 152
ID N80115 standard; DNA: 832 BP.

N80115;
AT 19-OCT-1990 (first entry)
DE Human preprolactin gene.
KW Prolactin; milk; contraceptive; dairy cows; lactation; as.
OS Homo sapiens. Location/Qualifiers
FH Key
FI Location/Qualifiers
FT 161..628
FT /tag= a
FT /product= preprolactin
FN US4725549-A.
PD 16-FEB-1988.
PR 21-MAR-1984; 592714.
PR 21-MAR-1984; US-592714.
PA (BREC) University of California.
PI Cooke NF, Baster JD;
DR WPI: 88-090571/10.
CC DNA coding for prolactin - obtd. by prepn. of reverse transcript
PT of mRNA coding for prolactin and inserting into a transfer vector.
PS Claim 12; Fig 3; 10pp; English.
CC The cDNA was prepared from RNA extracted from prolactin secreting
CC tumour-obtd. pituitary gland. The cDNA was inserted into a vector
CC into expression vectors for the prodn. of prolactin which can be used
CC admin. to dairy cows to increase milk yield. It can also be used
CC as a female contraceptive and to ensure adequate milk prodn. for
CC research and other purposes.
SQ Sequence 832 BP; 216 A; 233 C; 182 G; 201 T;
Query Match 1.38; Score 15; DB 1; Length 832;
Best Local Similarity 100.00; Pred. No. 8.47e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3 catgacatcaag 17
Cp 663 CATGACATCAAG 649

RESULT 153
ID T74333 standard; cDNA: 832 BP.
AT T74333;
AT 11-FEB-1998 (first entry)
DE cDNA encoding human substitution prolactin.
KW Prolactin antagonist; phosphorylation; hyperprolactinemia;
KW Prolactinoma; prostate cancer; tumour; T-lymphoma; infertility;
KW Prolactinoma; prostate cancer; tumour; T-lymphoma; infertility;
OS Homo sapiens.
CC Synthetic.

Key CDS Location/Qualifiers
 PT Vaccine; diagnosis: epitope; passive immunisation: as.
 PS Haemophilus influenzae.
 FT CDS
 FT 164..529
 FT /tag= a
 W08049322-A.
 PD 11-JUN-1987; 003433
 PD 21-JUN-1987; 003433
 PR 31-DEC-1986; US-948364.
 PR 02-MAR-1987; US-020849.
 PR 11-DEC-1987; US-132073.
 PR 02-MAR-1987; US-020849.
 PA Deich BA. Klotzsch G. Green B.
 PA WPI: 88-205305/29.
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PS Disclosure: Fig 14: 13pp; English.
 CC A pure antigenic peptide or protein related to an epitope
 CC of Haemophilus influenzae is claimed. Also claimed is a
 CC antigenic determinant of an Haemophilus influenzae outer membrane
 CC protein, the transformed cell, a subunit vaccine in a pharmaceutical
 CC carrier, a method of immunising humans and an assay for Haemophilus
 CC influenzae. 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 SQ Sequence 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 Query Match 1.38; Score 14; DB 2; Length 850;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 63 tggctgtgtgttt 76
 QY 749 TCTCTCTCTCTCTTT 762

RESULT 155
 ID T90883 standard; DNA: 860 BP.
 IC T90883:1998 (first entry)
 DE Monocyte chemotactic protein 5 cDNA.
 KW Monocyte chemotactic protein 5; MCP-5; human; macrophage;
 KW chemokine; inhibitor; antiinflammatory; atherosclerosis;
 KW chronic diseases; arthritis; angiogenesis; tumour; metastasis;
 KW WPI: 97-297357/27.
 OS Homo sapiens. Location/Qualifiers
 FE Key 66..372
 FE CDS 66..372
 FE /tag= a
 FE /note= Claim 3"

Key CDS Location/Qualifiers
 PT Vaccine; diagnosis: epitope; passive immunisation: as.
 PS Haemophilus influenzae.
 FT CDS
 FT 164..529
 FT /tag= a
 W08049322-A.
 PD 11-JUN-1987; 003433
 PD 21-JUN-1987; 003433
 PR 31-DEC-1986; US-948364.
 PR 02-MAR-1987; US-020849.
 PR 11-DEC-1987; US-132073.
 PR 02-MAR-1987; US-020849.
 PA Deich BA. Klotzsch G. Green B.
 PA WPI: 88-205305/29.
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PS Disclosure: Fig 14: 13pp; English.
 CC A pure antigenic peptide or protein related to an epitope
 CC of Haemophilus influenzae is claimed. Also claimed is a
 CC antigenic determinant of an Haemophilus influenzae outer membrane
 CC protein, the transformed cell, a subunit vaccine in a pharmaceutical
 CC carrier, a method of immunising humans and an assay for Haemophilus
 CC influenzae. 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 SQ Sequence 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 Query Match 1.38; Score 14; DB 2; Length 850;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 63 tggctgtgtgttt 76
 QY 749 TCTCTCTCTCTTT 762

RESULT 155
 ID T90883 standard; DNA: 860 BP.
 IC T90883:1998 (first entry)
 DE Monocyte chemotactic protein 5 cDNA.
 KW Monocyte chemotactic protein 5; MCP-5; human; macrophage;
 KW chemokine; inhibitor; antiinflammatory; atherosclerosis;
 KW chronic diseases; arthritis; angiogenesis; tumour; metastasis;
 KW WPI: 97-297357/27.
 OS Homo sapiens. Location/Qualifiers
 FE Key 66..372
 FE CDS 66..372
 FE /tag= a
 FE /note= Claim 3"

Key CDS Location/Qualifiers
 PT Vaccine; diagnosis: epitope; passive immunisation: as.
 PS Haemophilus influenzae.
 FT CDS
 FT 164..529
 FT /tag= a
 W08049322-A.
 PD 11-JUN-1987; 003433
 PD 21-JUN-1987; 003433
 PR 31-DEC-1986; US-948364.
 PR 02-MAR-1987; US-020849.
 PR 11-DEC-1987; US-132073.
 PR 02-MAR-1987; US-020849.
 PA Deich BA. Klotzsch G. Green B.
 PA WPI: 88-205305/29.
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PS Disclosure: Fig 14: 13pp; English.
 CC A pure antigenic peptide or protein related to an epitope
 CC of Haemophilus influenzae is claimed. Also claimed is a
 CC antigenic determinant of an Haemophilus influenzae outer membrane
 CC protein, the transformed cell, a subunit vaccine in a pharmaceutical
 CC carrier, a method of immunising humans and an assay for Haemophilus
 CC influenzae. 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 SQ Sequence 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 Query Match 1.38; Score 14; DB 2; Length 850;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 63 tggctgtgtgttt 76
 QY 749 TCTCTCTCTCTTT 762

RESULT 155
 ID T90883 standard; DNA: 860 BP.
 IC T90883:1998 (first entry)
 DE Monocyte chemotactic protein 5 cDNA.
 KW Monocyte chemotactic protein 5; MCP-5; human; macrophage;
 KW chemokine; inhibitor; antiinflammatory; atherosclerosis;
 KW chronic diseases; arthritis; angiogenesis; tumour; metastasis;
 KW WPI: 97-297357/27.
 OS Homo sapiens. Location/Qualifiers
 FE Key 66..372
 FE CDS 66..372
 FE /tag= a
 FE /note= Claim 3"

Key CDS Location/Qualifiers
 PT Vaccine; diagnosis: epitope; passive immunisation: as.
 PS Haemophilus influenzae.
 FT CDS
 FT 164..529
 FT /tag= a
 W08049322-A.
 PD 11-JUN-1987; 003433
 PD 21-JUN-1987; 003433
 PR 31-DEC-1986; US-948364.
 PR 02-MAR-1987; US-020849.
 PR 11-DEC-1987; US-132073.
 PR 02-MAR-1987; US-020849.
 PA Deich BA. Klotzsch G. Green B.
 PA WPI: 88-205305/29.
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 DR Pure peptide related to epitope of Haemophilus influenzae -
 DR P-PSDB: P00667.
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PT Passive immunisation and for producing antibodies -
 PS Disclosure: Fig 14: 13pp; English.
 CC A pure antigenic peptide or protein related to an epitope
 CC of Haemophilus influenzae is claimed. Also claimed is a
 CC antigenic determinant of an Haemophilus influenzae outer membrane
 CC protein, the transformed cell, a subunit vaccine in a pharmaceutical
 CC carrier, a method of immunising humans and an assay for Haemophilus
 CC influenzae. 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 SQ Sequence 850 BP; 254 A; 116 C; 190 G; 288 T; 2 Others;
 Query Match 1.38; Score 14; DB 2; Length 850;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 63 tggctgtgtgttt 76
 QY 749 TCTCTCTCTCTTT 762

RESULT 155
 ID T90883 standard; DNA: 860 BP.
 IC T90883:1998 (first entry)
 DE Monocyte chemotactic protein 5 cDNA.
 KW Monocyte chemotactic protein 5; MCP-5; human; macrophage;
 KW chemokine; inhibitor; antiinflammatory; atherosclerosis;
 KW chronic diseases; arthritis; angiogenesis; tumour; metastasis;
 KW WPI: 97-297357/27.
 OS Homo sapiens. Location/Qualifiers
 FE Key 66..372
 FE CDS 66..372
 FE /tag= a
 FE /note= Claim 3"

US-08-887-977-9 - rna

Seq	Sequence	967 BP	358 A	188 C	202 G	219 T
	Query Match					
	Best Local Similarity	100.0%				
	Matches	14	Conservative	0	Mismatches	0
						Indels 0; Caps 0;
DB	206 caaaatcatcatgaa	219				
CP	671 caaaatcatcatgaa	658				

RESULT 168

ID: Q04289 standard; DNA: 983 BP.

AC Q04289-1990 (first entry)

DC Sequence encoding ovine trophoblast protein-1 (ovp-1).

DE Bovine trophoblast protein-1; btp-1; fertility; da

KW Sus scrofa.

PN 2P-187063A.

PP 23-OCT-1989; 119642.

PR 26-OCT-1988; US-762870.

PX (CMOR) Univ of Missouri.

PI Roberts RM, Mahavak K;

PS P-PSBR; 80450/15.

DR P-PSBR; 80450.

PT Recombinant bovine trophoblast protein-1 -
used for enhancing fertility or treating viral diseases in
mammals, esp. cattle.

CC The btp-1 produced from the gene may be used to promote fertility
or treat viral disease in cattle. The gene may also be used to
provide transgenic animals with enhanced fertility, or in
concomitant with other genetic treatments to enhance fertility,
prophylactic and therapeutic treatment of other mammals.

QC Sequence 983 BP; 283 A; 280 C; 221 G; 239 T;

Query Match

Best Local Similarity 100.0%; Score 14; DB 1; Length 983;
Matches 14; Conservative 0; Pred. No. 3,476+02;

Dy 311 ccctgctgcattac 324
|||||
Qy 909 cccgtcgtcttacc 922

RESULT 169

ID: Q47200 standard; DNA: 1002 BP.

AC Q47200-1994 (first entry)

DC 26-JAN-1994

DE RNA polyoma-virus 'holoeyme delta' subunit holo gene.

KW

PN

PP

PR

PX

PI

PS

DR

PT

CC

QC

Query Match

Best Local Similarity 100.0%; Score 14; DB 1; Length 983;
Matches 14; Conservative 0; Pred. No. 3,476+02;

Dy 311 ccctgctgcattac 324
|||||
Qy 909 cccgtcgtcttacc 922

Page 180

PF C/Gaskell NF;
DR WPI; 97-470038/43.
DR P-FDB; W36071.
DR P-Isolated theta, chi, psi, delta and delta' protein subunits of E.
CC coli. The delta and delta' subunits were made man-made enzymes of 5
CC amino acids useful for long chain PCR
PF Clam 35; Column 19-20; 65pp; English.
CC This is the nucleotide sequence encoding the delta' subunit of the
CC E. coli polymerase III holoenzyme. The gene was isolated from
CC a lambda phage library using probes 19519-19559.
CC The E. coli polymerase III subunits (theta, chi, psi, delta and delta'
CC 19519-19559 respectively) are used as man-made enzymes
CC for cloning and amplification. They are available at the China PCR
CC Centre. Sequence ID: 19567 C; 275 G; 242 T.
CC Sequence 1002 BP; 218 A; 247 C; 247 C; 242 T.

Query Match 138; Score 14; DP 36; Length 1002;
Best Local Similarity 100.0%; Pred. No; 3 47+02;
Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 379 gttcttcggcagac 392
|||||
Cp 320 cgaattcgtaaac 307

RESULT 171
ID Q34508 standard; cdna to rRNA; 1033 BP.

DE	Key	Location/Qualifiers
DE	cds	41-388
DE	cds	41-388 a
DE	misc_Rna	119-385
DE	misc_Rna	119-385 b
DE	misc_Rna	119-385 c
DE	misc_Rna	119-385 d
DE	misc_Rna	119-385 e
DE	misc_Rna	119-385 f
DE	misc_Rna	119-385 g
DE	misc_Rna	119-385 h
DE	misc_Rna	119-385 i
DE	misc_Rna	119-385 j
DE	misc_Rna	119-385 k
DE	misc_Rna	119-385 l
DE	misc_Rna	119-385 m
DE	misc_Rna	119-385 n
DE	misc_Rna	119-385 o
DE	misc_Rna	119-385 p
DE	misc_Rna	119-385 q
DE	misc_Rna	119-385 r
DE	misc_Rna	119-385 s
DE	misc_Rna	119-385 t
DE	misc_Rna	119-385 u
DE	misc_Rna	119-385 v
DE	misc_Rna	119-385 w
DE	misc_Rna	119-385 x
DE	misc_Rna	119-385 y
DE	misc_Rna	119-385 z
DE	misc_Rna	119-385 aa
DE	misc_Rna	119-385 ab
DE	misc_Rna	119-385 ac
DE	misc_Rna	119-385 ad
DE	misc_Rna	119-385 ae
DE	misc_Rna	119-385 af
DE	misc_Rna	119-385 ag
DE	misc_Rna	119-385 ah
DE	misc_Rna	119-385 ai
DE	misc_Rna	119-385 aj
DE	misc_Rna	119-385 ak
DE	misc_Rna	119-385 al
DE	misc_Rna	119-385 am
DE	misc_Rna	119-385 an
DE	misc_Rna	119-385 ao
DE	misc_Rna	119-385 ap
DE	misc_Rna	119-385 aq
DE	misc_Rna	119-385 ar
DE	misc_Rna	119-385 as
DE	misc_Rna	119-385 at
DE	misc_Rna	119-385 au
DE	misc_Rna	119-385 av
DE	misc_Rna	119-385 aw
DE	misc_Rna	119-385 ax
DE	misc_Rna	119-385 ay
DE	misc_Rna	119-385 az
DE	misc_Rna	119-385 ba
DE	misc_Rna	119-385 bb
DE	misc_Rna	119-385 bc
DE	misc_Rna	119-385 bd
DE	misc_Rna	119-385 be
DE	misc_Rna	119-385 bf
DE	misc_Rna	119-385 bg
DE	misc_Rna	119-385 bh
DE	misc_Rna	119-385 bi
DE	misc_Rna	119-385 bj
DE	misc_Rna	119-385 bk
DE	misc_Rna	119-385 bl
DE	misc_Rna	119-385 bm
DE	misc_Rna	119-385 bn
DE	misc_Rna	119-385 bo
DE	misc_Rna	119-385 bp
DE	misc_Rna	119-385 bq
DE	misc_Rna	119-385 br
DE	misc_Rna	119-385 bs
DE	misc_Rna	119-385 bt
DE	misc_Rna	119-385 bu
DE	misc_Rna	119-385 bv
DE	misc_Rna	119-385 bw
DE	misc_Rna	119-385 bx
DE	misc_Rna	119-385 by
DE	misc_Rna	119-385 bz
DE	misc_Rna	119-385 ca
DE	misc_Rna	119-385 cb
DE	misc_Rna	119-385 cc
DE	misc_Rna	119-385 cd
DE	misc_Rna	119-385 ce
DE	misc_Rna	119-385 cf
DE	misc_Rna	119-385 cg
DE	misc_Rna	119-385 ch
DE	misc_Rna	119-385 ci
DE	misc_Rna	119-385 cj
DE	misc_Rna	119-385 ck
DE	misc_Rna	119-385 cl
DE	misc_Rna	119-385 cm
DE	misc_Rna	119-385 cn
DE	misc_Rna	119-385 co
DE	misc_Rna	119-385 cp
DE	misc_Rna	119-385 cq
DE	misc_Rna	119-385 cr
DE	misc_Rna	119-385 cs
DE	misc_Rna	119-385 ct
DE	misc_Rna	119-385 cu
DE	misc_Rna	119-385 cv
DE	misc_Rna	119-385 cw
DE	misc_Rna	119-385 cx
DE	misc_Rna	119-385 cy
DE	misc_Rna	119-385 cz
DE	misc_Rna	119-385 da
DE	misc_Rna	119-385 db
DE	misc_Rna	119-385 dc
DE	misc_Rna	119-385 dd
DE	misc_Rna	119-385 de
DE	misc_Rna	119-385 df
DE	misc_Rna	119-385 dg
DE	misc_Rna	119-385 dh
DE	misc_Rna	119-385 di
DE	misc_Rna	119-385 dj
DE	misc_Rna	119-385 dk
DE	misc_Rna	119-385 dl
DE	misc_Rna	119-385 dm
DE	misc_Rna	119-385 dn
DE	misc_Rna	119-385 do
DE	misc_Rna	119-385 dp
DE	misc_Rna	119-385 dq

Db 873 gcagatacacaat 887
QY 1063 GCAGATACACAAT 1077

RESULT 176
ID Q29138 standard; mRNA; 1040 BP.
AC Q29138;
DE P1-MAR-1993 (first entry)
KW Protein ID: P1B; Neisseria gonorrhoeae; P1A; outer membrane; porin;
KW hydrophobic; lipid; outer membrane; antigen; immunogen; vaccine;
KW gonorrhoea; infection; as.
PM 86
PR 86
PS 86
FT Key Location/Qualifiers
FT CDS 81..1028
FT /tag= a
FT signal_peptide 81..134
FT /tag= b
FT mat_peptide 135..1025
FT /tag= c
FT -35_signal 31..36
FT /tag= d
FT -10_signal /tag= e
FT rbs 69..73
FT /tag= f
FT misc_feature 207..210
FT /tag= g
FT misc_feature 246..249
FT /tag= h
FT cds
PM W09216225.A
PM 13-MAR-1993; 002006.
PP 13-MAR-1993; 002006.
PR 14-MAR-1991; US-669492.
PA (UNIC-) UNIV NORTH CAROLINA.
DI Carbonetti, M.H., Sparling, P.F.
DR P-PSDB; 827483.
FR Fragments of Neisseria gonorrhoeae protein IA or IB - used as
FR vaccines for preventing gonorrhoeal infection, and for diagnosis
FR of gonorrhoea. Fig 3; 59pp; English.
CC The gene for protein IA, a major outer membrane protein of N. gonorrhoeae
CC which act as a porin. Porins are believed to act in cells by
CC channeling low molecular weight substances across the hydrophobic
CC lipid bilayer. The gene sequences of P1A and P1B are identical.
CC P1B allows the production of antibodies to the gene sequences of P1A and
CC techniques. These fragments can be used as immunogens in vaccine
CC compositions for the prevention of gonorrhoea, and also on immunosays
CC for diagnosis of infection. See also Q29139.
CC
SQ Sequence 1040 BP; 277 A; 282 C; 257 G; 224 T;
Query Match 1.3%; Score 15; DB 5; Length 1040;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 873 gcagatacacaat 887
QY 1063 GCAGATACACAAT 1077

RESULT 177
ID N90048 standard; DNA; 1040 BP.
AC N90048;
DE P1-MAR-1989 (first entry)
KW Neisseria gonorrhoeae (FA19); Protein IA DNA
KW Neisseria gonorrhoeae (NG); NG strain FA19; DNA; Protein IA.
PM 86
PR 86
PS 86
FT Key Location/Qualifiers
FT CDS 81..1028
FT /tag= a
FT -35_signal 31..36
FT /tag= b
FT -10_signal 50..55
FT /tag= c
FT rbs 69..73
FT /tag= d
FT misc_signal 81..134
FT /tag= e
FT misc_feature 207..210
FT /tag= f
FT misc_feature 246..249
FT /tag= g
FT misc_feature 246..249
FT /tag= h
FT cds
PM W0904873.A
PM 01-JUN-1989; 004225.
PP 01-JUN-1989; 004225.
PR 21-JUN-1989; US-672758.
PA (UNIC) University of North Carolina.
DI Carbonetti, M.H., Sparling, P.F.
DR WPI; 89-178391/24.
FR Neisseria gonorrhoeae Protein I gene sequences
FR - Used for detection of N. gonorrhoeae infection and
FR for producing vaccines for prevention or treatment of infection
FR (disclosure; Fig 3; 70pp; English).
CC The gene for protein IA of NG strain FA19
CC also encodes a protein of 1025 amino acids. The gene
CC sequence is 1040 bp. The gene contains 10 exons and 9 introns.
CC peptide. (f) is a SalI site, and (g) is a TagI site (see N90035 -
CC N90048).
SQ Sequence 1040 BP; 277 A; 286 C; 253 G; 224 T;
Query Match 1.3%; Score 15; DB 1; Length 1040;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 873 gcagatacacaat 887
QY 1063 GCAGATACACAAT 1077

RESULT 178
ID T03883 standard; DNA; 1070 BP.
AC T03883;
DE 25-MAY-1996 (first entry) kinase family ligand, EFL-2 encoding DNA.
DE EFL-2; EFL-2; Eph transmembrane tyrosine kinase family ligand;
KW neurological disorder; identification; diagnosis; as.
OS Homo sapiens.
PM 86
PR 86
PS 86
FT Key Location/Qualifiers
FT CDS 47..750
FT /tag= a
FT /product= EFL-2
FT W09527060.A2
PM 04-APR-1996; 004208.
PP 04-APR-1996; 004208.
PR 12-APR-1994; US-229075.
PR 12-APR-1994; US-229075.
PR 31-SEP-1994; US-229075.
PR 31-SEP-1994; US-229075.
PA (REGZ) REGENERON PHARM INC.
PI Aldrich, R.H., Davis, S., Gale, N., Goldfarb, M., Maisonspierre, D.C.;
PI Tarcopoules, G.D.
PM P-PSDB; 5-386257/46.
DR P-PSDB; 5-386257/46.
PT Ligands which bind Eph family receptors - used in the diagnosis of
PT neurological disorders
PS Disclosure; Fig 3; 59pp; English.
CC The DNA encodes a Eph transmembrane tyrosine kinase family ligand
CC designated EFL-2. The ligand is a member of the EFL family of
CC EFL-1, -2, -3, EFL and EFL receptors. The ligands are useful in
CC promoting a differential function and/or influencing the phenotype,
CC such as growth and/or proliferation, of receptor bearing cells. They
CC can be used for diagnosis, and treatment of neurological disorders.
SQ Sequence 1070 BP; 339 A; 341 C; 334 G; 189 T;
Query Match 1.3%; Score 14; DB 18; Length 1070;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 873 gcagatacacaat 887
QY 1063 GCAGATACACAAT 1077

RESULT 180
ID T27350 standard; DNA; 1087 BP.
AC T27350;
DE 13-NOV-1996 (first entry)
DE Meloidogyne resistant tomato acid phosphatase 1(1) (Aps1(1)) gene.
KW Acid phosphatase 1; Aps1 gene; Meloidogyne; tomato; detection;
KW determination; genotype; resistance; sensitivity; amplification;
KW PCR; gel electrophoresis; agricultural; hybridization; da.
OS Lycopersicon esculentum.

Query Match 1.3%; Score 15; DB 1; Length 1080;
Best Local Similarity 100.0%; Pred. No. 8.74e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 41 cacaagaccacacac 55
Cp 759 CACAGACCACACAC 745

RESULT 180
ID T27350 standard; DNA; 1087 BP.
AC T27350;
DE 13-NOV-1996 (first entry)
DE Meloidogyne resistant tomato acid phosphatase 1(1) (Aps1(1)) gene.
KW Acid phosphatase 1; Aps1 gene; Meloidogyne; tomato; detection;
KW determination; genotype; resistance; sensitivity; amplification;
KW PCR; gel electrophoresis; agricultural; hybridization; da.
OS Lycopersicon esculentum.

PI (SNOW) SNOW BRAND MILK PRODUCTS.
 DR Morinaga T, Yasuda N, Higashio K, Tamaoki T;
 DR WPI: 90-161490/48.
 PR P-PSD8: W47410.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 PT protein expression in animal cells.
 PS Claim 2, Page 21; 36pp; Japanese.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 CC gene. The enhancer is obtained from human hepatoma cells and promotes
 CC Large amounts of the protein can be produced. See also Q07005.
 SQ Sequence 1091 BP; 300 A; 302 C; 253 G; 236 T;
 Query Match 1.38; Score 14; DB 1; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 106 cccagctccagctga 119
 QY 1062 cccagctccagctga 1075

RESULT 182
 ID Q77819 standard; cDNA: 1093 BP.
 AC Q77819;
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"

PI Banada K, Hara R, Matsuki Y, Shindo S, Yoshimura H;
 DR WPI: 94-35870/39.
 PR P-PSD8: W47410.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 CC gene. The enhancer is obtained from human hepatoma cells and promotes
 CC Large amounts of the protein can be produced. See also Q07005.
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;
 Query Match 1.38; Score 14; DB 1; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 106 cccagctccagctga 119
 QY 1062 cccagctccagctga 1075

DT 05-JUN-1998 (first entry)
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"

PI Banada K, Hara R, Matsuki Y, Shindo S, Yoshimura H;
 DR WPI: 94-35870/39.
 PR P-PSD8: W47410.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 CC gene. The enhancer is obtained from human hepatoma cells and promotes
 CC Large amounts of the protein can be produced. See also Q07005.
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;
 Query Match 1.38; Score 14; DB 1; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 106 cccagctccagctga 119
 QY 1062 cccagctccagctga 1075

RESULT 185
 ID V15793 standard; cDNA to mRNA: 1095 BP.
 AC V15793;
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"

FT Key
 FT Location/Qualifiers
 FT 1..184
 FT /tag= "a"
 FT 1008-1087
 FT /tag= "b"
 FT /number= 2

PI Banada K, Hara R, Matsuki Y, Shindo S, Yoshimura H;
 DR WPI: 94-35870/39.
 PR P-PSD8: W47410.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 CC gene. The enhancer is obtained from human hepatoma cells and promotes
 CC Large amounts of the protein can be produced. See also Q07005.
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;
 Query Match 1.38; Score 14; DB 1; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 106 cccagctccagctga 119
 QY 1062 cccagctccagctga 1075

RESULT 181
 ID Q06624 standard; DNA: 1091 BP.
 AC Q06624;
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"

DT 05-JUN-1998 (first entry)
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"

PI Banada K, Hara R, Matsuki Y, Shindo S, Yoshimura H;
 DR WPI: 94-35870/39.
 PR P-PSD8: W47410.
 CC The present sequence is a cDNA for wild type human L-asparaginase homologue.
 CC The enhancer is obtained from human hepatoma cells and promotes
 CC gene. The enhancer is obtained from human hepatoma cells and promotes
 CC Large amounts of the protein can be produced. See also Q07005.
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;
 Query Match 1.38; Score 14; DB 1; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 106 cccagctccagctga 119
 QY 1062 cccagctccagctga 1075

RESULT 184
 ID V15793 standard; cDNA to mRNA: 1095 BP.
 AC V15793;
 DE cDNA for wild type human L-asparaginase homologue.
 DE Leukemia; lymphoma; human; ds.
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;
 KW Liver cancer; cirrhosis; as.
 OS Homo sapiens.
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1095
 FT /note= "stop codon not given"


```

cde      1042..1132
/ctg=b   525..600
/ctg=a   c

wt0305164.A.
18-MAR-1993.  CB-018759.
02-BER-1991.  CB-018759.
(DTLE.) UNIV LEICESTER.
PA Drapes, J. 000984/12.
PR P-RSD; R325B.
DR DNA comprising marker gene and callus-specific promoter - for
  transformation of plants e.g. asparagus following wounding and/or
  Claim 3 formation.
CC This sequence represents the A. officinalis AOP1 promoter which was
  generated by inverse polymerase chain reaction (IPCR) (see also
  CB018751-61). The Aop1 gene is a wound or pathogen induced gene. It
  is 1731 bp long, contains 10 exons and 11 introns. It is highly
  CC quickly upregulated and highly active in wounded plant material both
  in wounded seedlings and in tissue cultured explants. Its high
  CC expression persists for several days post-wounding. The majority of
  CC the Aop1 promoter is used to drive a marker gene in transgenic
  CC plant only a minimal amount of unwanted marker gene product will be
  CC produced.
  SQ produced. 1132 BP; 353 A; 233 C; 195 G; 351 T;

      1.38; Score 14; DB 6; Length 1132;
      Best Local Similarity 100.0%; Pred. No. 3.47e+02;
      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
      0

Db      730 tctaaagccgca 743
      |||||
Qy      712 tctaaagccgca 725

RESULT 192
ID G13697 standard; cDNA; 1152 BP.
CC 17-MAR-1992 (first entry)
DE Rat1a encoding capsid region of cyno-HAV isolate CT-55/JN-55.
KW Repatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque; aa.
PR Cynomolgus monkey hepatitis A virus, isolate CT-55/JN-55.
PA KPM
PR P-RSD; R325B.
DR DNA comprising marker gene and callus-specific promoter - for
  transformation of plants e.g. asparagus following wounding and/or
  Claim 3 formation.
CC This sequence represents the A. officinalis AOP1 promoter which was
  generated by inverse polymerase chain reaction (IPCR) (see also
  CB018751-61). The Aop1 gene is a wound or pathogen induced gene. It
  is 1731 bp long, contains 10 exons and 11 introns. It is highly
  CC quickly upregulated and highly active in wounded plant material both
  in wounded seedlings and in tissue cultured explants. Its high
  CC expression persists for several days post-wounding. The majority of
  CC the Aop1 promoter is used to drive a marker gene in transgenic
  CC plant only a minimal amount of unwanted marker gene product will be
  CC produced.
  SQ produced. 1132 BP; 353 A; 233 C; 195 G; 351 T;

      1.128 b
/ctg=a   b

```

```

FT      /tag= b
          296 *301
FT      /*tag= c
          320..325
FT      /tag= d
          320..325
PD      WQ9308286-A.
PP      29-APR-1993.
PR      21-OCT-1992; US781355.
PT      22-OCT-1991; US781355.
PQ      22-OCT-1991; US781355.
PA      Barbour AG, Norris RM.
PB      WPI; 93-152478/18.
PC      P-PDB: R36675.
PD      DNA segment encoding 30 kD Borrelia burgdorferi antigenic
PE      protein subunit B, which has been used to detect Borrelia
PF      burgdorferi-associated diseases*, e.g. Lyme disease
PG      Claim 1; Fig 9; 8pp; English.
PH      The 30 kD protein was identified in low-passage virulent strains of
PI      the B. burgdorferi strain N40. This protein was partially sequenced and an invariant
PJ      sequence was determined. An oligonucleotide probe was synthesized and used to identify DNA
PK      containing the 30 kD protein in a low-passage virulent B. burgdorferi
PL      RNA library.
PM      Sequence 1161 BP; 519 A; 137 C; 169 G; 337 T;
PN      Query Match
PO      Best Local Similarity 100.0%; Pred. No. 3.47e+02;
PP      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
PQ      Db 551 ccnaaatatcaca 564
PY      Qy 531 CCMAAAATACACA 544
PS      RESULT 194
PT      AC T79695 standard; DNA; 1164 BP.
PU      AC T79695
PV      AC T79695
PW      DT 03-FEB-1998 (first entry)
PX      DE BRCA1 cancer susceptibility gene exon 16 and intron boundaries.
PY      KW BRCA1 cancer susceptibility gene; breast cancer; ovarian cancer;
PZ      melanoma; colorectal cancer; prostate cancer; colorectal cancer; ocular melanoma;
QA      leukemia; lymphoma; esophagus; stomach; colon; rectum; bladder;
QB      esophagus; stomach; colon; rectum; bladder; kidney; testis;
QC      esophagus; stomach; colon; rectum; bladder; kidney; testis;
QD      Homo sapiens.
QE      Location/Qualifiers
QF      Key 1..480
QG      Intron 1..480
QH      Intron 1..480
QI      Intron 1..480
QJ      Intron 1..480
QK      exon
QL      exon
QM      exon
QN      exon
QO      exon
QP      exon
QJ      exon

```


Page 216

50 clover yellow ve

Db 921 aaaaaaacaaaggtg 934

50 clover yellow ve

OS Mycobacterium tuberculosis
 PH CDS
 FT 88 1122
 FT /tag= a
 FT /product= Antigen_Tbaal7
 PD W0970420-22
 PD 134461994
 PD 134461994
 PD 30-AUG-1996: US-680573
 PD 12-JUL-1996: US-680573
 PD 01-SEP-1995: US-521453
 PD 22-MAR-1995: US-521453
 PD 22-MAR-1995: US-521453
 PD 05-JUN-1996: US-658800
 PD (CORI-) CORIXA CORP.
 PD Campos-neto A, Dillon DC, Houghton R, Read SG, Skealy YAW,
 PD WPI: 97-192904/21
 PD P-PSDB: W33357
 PD Query Match 1.38; Score 14; DB 34; Length 1362;
 PD Best Local Similarity 100.04; Pred. No. 3.47e-02; Indels 0; Gaps 0;
 PD Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1186 caggtgaagctgc 1209
 CY 597 CAGGAGGAGGCTTCA 610

RESULT 219
 ID W07719 standard: DNA: 1377 BP.
 DT 19-MAY-1998 (first entry)
 DE CP4-EPSPS gene.
 KW 1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;
 KW fruit ripening; ethylene production; glyphosate resistance;
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; EPSPS; CP4; ss.
 OS Synthetic.

PR 31-AUG-1990: US-576537.
 PR 28-AUG-1991: US-748611.
 PR 28-AUG-1991: US-748611.
 PR (MONS) MONSANTO CO.
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-297418/27.
 DR New isolated 5-enol-pyruvyl:shikimate-3-phosphate synthase gene -
 PT glyphosate herbicidal plants to produce plants which are tolerant to
 PS Glyphosate herbicidal plants to produce plants which are tolerant to
 PS Claim 3: Column 75-76; 154pp; English.
 CC 777314 encodes a class II 5-enolpyruvylshikimate-3-phosphate synthase
 CC (EPSPS) enzyme, derived from Agrobacterium sp. The sequence was altered
 CC to produce a mutant enzyme that may effect expression or
 CC stability in plant cell and tissues of a plant. The mutant enzyme
 CC or more and A+T rich regions that could function as polyadenylation
 CC sites or potential RNA destabilization region. Class II EPSPS enzymes
 CC are tolerant to glyphosate herbicides. EPSPS and sequences encoding it
 CC are tolerant to glyphosate herbicides. EPSPS and sequences encoding it
 CC plants such as corn, wheat, rice, barley, soybean, cotton, sugarbeet,
 CC oilseed rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa,
 CC poplar, pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf
 CC grasses.
 SQ Sequence 1377 BP: 278 A; 345 C; 355 G; 399 T;
 Query Match 1.38; Score 14; DB 32; Length 1377;
 Best Local Similarity 100.04; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 95 ccacaggtccttca 108
 CP 978 CCACAGGAGGCTTCA 965

RESULT 221
 ID Q23531 standard: DNA: 1377 BP.
 AC Q23531 999 (first entry)
 DE Synthetic Class II EPSPS gene.
 KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
 OS Synthetic.
 PD WPI: 97-114136/21
 PD 19-MAR-1992
 PD 28-AUG-1991: US-6148.
 PD 31-AUG-1990: US-576537.
 PD (MONS) MONSANTO CO.
 PD Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-114136/21
 PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
 PT synthase - for producing plants and bacteria tolerant to
 PT glyphosate herbicides

PH Key
 PH CDS
 FT 3 1370
 FT /tag= a
 FT /product= C74-EPSPS
 PD US5702933-A.
 PD 30-DEC-1997
 PD 30-DEC-1997: US-553843
 PD 30-DEC-1997: US-553843
 PD 26-DEC-1990: US-632440
 PD 06-NOV-1995: US-553943
 PD (MONS) MONSANTO CO.
 PD WPI: 97-074139/20
 PD WPI: 98-074139/20
 PD P-PSDB: W39426
 PD Production of plants with delayed ripening - using DNA encoding
 PD 1-aminocyclopropane-1-carboxylic acid deaminase
 PD 1-aminocyclopropane-1-carboxylic acid deaminase
 PD This sequence encodes the synthetic CP4 5-enolpyruvyl-3-shikimate
 CC phosphate synthase (EPSPS) gene which is capable of conferring
 CC resistance to glyphosate and is used in a novel method for producing
 CC fruit-bearing plants with delayed ripening. The method involves the
 CC introduction of a DNA encoding the synthetic CP4 5-enolpyruvyl-3-shikimate
 CC gene in a plant at a level sufficient to produce ethylene production in
 CC the fruit.
 SQ Sequence 1377 BP: 278 A; 344 C; 355 G; 400 T;
 Query Match 1.38; Score 14; DB 39; Length 1377;
 Best Local Similarity 100.04; Pred. No. 3.47e-02; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 95 ccacaggtccttca 108
 CP 978 CCACAGGAGGCTTCA 965

RESULT 220
 ID 777314 standard: DNA: 1377 BP.
 AC 777314
 DT 02-OCT-1997 (first entry)
 DE EPSPS DNA having sequence altered for increased expression in plants.
 DE This sequence encodes the synthetic CP4 5-enolpyruvyl-3-shikimate
 KW glyphosate resistant; transgenic plant; herbicide; shikimate;
 KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
 KW tobacco; da.
 OS Synthetic.
 Location/Qualifiers
 FT CDS
 FT 3 1370
 FT /tag= a
 PD US5633435-A.
 PD 31-AUG-1990: 576537.
 PD 31-AUG-1990: 576537.

PS Disclosure: Fig 8: 14pp; English.
 CC The sequence is that of a synthetic Class II 5'-enolpyruvylshikimate-3
 CC phosphate synthase enzyme (EPSPS) gene. It can be used to create
 CC glyphosate resistant plants or seeds by transforming plant cells.
 CC The resulting plants or seeds are tolerant to glyphosate herbicides
 CC weeds. The crops selected for are e.g. corn, wheat, rice, oilseed
 CC rape, tobacco and alfalfa. This provides a cost effective,
 CC environmentally compatible weed control device.
 SQ Sequence 1377 BP: 278 A; 345 C; 355 G; 399 T;
 Query Match 1.38; Score 14; DB 3; Length 1377;
 Best Local Similarity 100.04; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 95 ccacaggtccttca 108
 CP 978 CCACAGGAGGCTTCA 965

RESULT 222
 ID 793791 standard: DNA: 1377 BP.
 DC 17-791-1998 (first entry)
 DE Synthetic CP4 Class II EPSP synthase (EPSPS) gene.
 KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
 KW shikimate acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
 KW 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
 KW Synthetic.
 OS Synthetic.
 PD US5627061-A.
 PD 06-MAY-1997.
 PD 07-JUN-1995; 476008
 PD 31-AUG-1990: US-576537.
 PD 31-AUG-1990: US-576537.
 PD 28-AUG-1991: US-749611.
 PD (MONS) MONSANTO CO.
 PD Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-114136/21
 PT Production of glyphosate-herbicide tolerant plants - using DNA
 PT encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
 PT enzyme(s)
 OS Synthetic.
 PD WPI: 97-114136/21
 PD 28-AUG-1991: US-6148.
 PD 31-AUG-1990: US-576537.
 PD (MONS) MONSANTO CO.
 PD Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-114136/21
 PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
 PT synthase - for producing plants and bacteria tolerant to
 PT glyphosate herbicides

PR 31-AUG-1990: US-576537.
 PR 28-AUG-1991: US-748611.
 PR 28-AUG-1991: US-748611.
 PR (MONS) MONSANTO CO.
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-297418/27.
 DR New isolated 5-enol-pyruvyl:shikimate-3-phosphate synthase gene -
 PT glyphosate herbicidal plants to produce plants which are tolerant to
 PS Glyphosate herbicidal plants to produce plants which are tolerant to
 PS Claim 3: Column 75-76; 154pp; English.
 CC 777314 encodes a class II 5-enolpyruvylshikimate-3-phosphate synthase
 CC (EPSPS) enzyme, derived from Agrobacterium sp. The sequence was altered
 CC to produce a mutant enzyme that may effect expression or
 CC stability in plant cell and tissues of a plant. The mutant enzyme
 CC or more and A+T rich regions that could function as polyadenylation
 CC sites or potential RNA destabilization region. Class II EPSPS enzymes
 CC are tolerant to glyphosate herbicides. EPSPS and sequences encoding it
 CC are tolerant to glyphosate herbicides. EPSPS and sequences encoding it
 CC plants such as corn, wheat, rice, barley, soybean, cotton, sugarbeet,
 CC oilseed rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa,
 CC poplar, pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf
 CC grasses.
 SQ Sequence 1377 BP: 278 A; 345 C; 355 G; 399 T;
 Query Match 1.38; Score 14; DB 32; Length 1377;
 Best Local Similarity 100.04; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 95 ccacaggtccttca 108
 CP 978 CCACAGGAGGCTTCA 965

RESULT 221
 ID Q23531 standard: DNA: 1377 BP.
 AC Q23531 999 (first entry)
 DE Synthetic Class II EPSPS gene.
 KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
 OS Synthetic.
 PD WPI: 97-114136/21
 PD 19-MAR-1992
 PD 28-AUG-1991: US-6148.
 PD 31-AUG-1990: US-576537.
 PD (MONS) MONSANTO CO.
 PD Barry GF, Kishore GM, Padgett SR, Stallings WC;
 DT WPI: 97-114136/21
 PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
 PT synthase - for producing plants and bacteria tolerant to
 PT glyphosate herbicides

CC (either as a fused protein or as a cell extract) and analysing the
 CC the amino acid sequence of the protein. The amino acid sequence
 CC or not rendered carcinogenic or "unsafe" if it is not detoxified or is
 CC metabolised to a carcinogenic cpd.
 SQ Sequence 1473 BP; 424 A; 335 C; 313 G; 401 T;
 Query Match 1 39; Score 15; DB 22; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951
 RESULT 235
 ID T17394 standard; cDNA; 1473 BP.
 AC T17394;
 DT 01-AUG-1996 (first entry)
 DE Human derived cytochrome P450C28 cDNA.
 DE Human derived cytochrome P450C28 cDNA.
 DE Human derived cytochrome P450C28 cDNA.
 KW Liver; Yeast; expression vector; WADPH-P450 reductase; ADH gene promoter;
 KW evaluation; safety; fusion protein; metabolite; detoxification;
 KW carcinogenic; ds.
 OS Homo sapiens.
 FT cd5 Location/Qualifiers
 FT 1..1473 /tag= a
 FT 108027198-A.
 PD 10-JAN-1996. 164185.
 PR 10-JAN-1996: JP-164185.
 PR 15-JUL-1994: JP-164185.
 PA (SUNO) SUMITOMO CHEM CO LTD.
 DR WPI: 96-13639/14.
 DR PUBMED 814651.
 DR Antisense RNA inhibiting human derived cytochrome P450C28 - allows
 PT specific detection of cytochrome P450 species in humans
 PT Example 1; Pages 10-12; 13pp; Japanese.
 CC The present sequence encodes the human derived cytochrome (HDC)
 CC cytochrome P450C28. The gene encodes a protein of 290 amino
 CC acids. The protein is a member of the cytochrome P450 family.
 CC were transfected with an expression vector. The resulting HDC
 CC cultured and then disrupted to give a microsomal fraction. The
 CC HDC was purified from the fraction, and used to immunise and
 CC sensitise a mammal. Blood was drawn from the mammal, and an
 CC antiserum was prepared. The antiserum was used to detect the
 CC P450C28, partic. at serum dilution rates of 1:10000, and is
 CC substantially without cross reaction to other HDC P450 spp.
 SQ Sequence 1473 BP; 425 A; 335 C; 313 G; 400 T;
 Query Match 1 39; Score 15; DB 20; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951
 RESULT 236
 ID T28330 standard; cDNA; 1473 BP.
 AC T28330;
 DT 11-OCT-1996 (first entry)
 DE Human cytochrome P450 molecular species 2C8 variant 1; gene.
 DE Human cytochrome P450 molecular species 2C8 variant 1; gene.
 DE Human cytochrome P450 molecular species 2C8 variant 1; gene.
 KW Liver; Yeast; expression vector; WADPH-P450 reductase; ADH gene promoter;
 KW evaluation; safety; fusion protein; metabolite; detoxification;
 KW carcinogenic; ds.
 OS Homo sapiens.
 FT cd5 Location/Qualifiers
 FT 1..1473 /tag= a
 FT 108027198-A.
 PD 10-JAN-1996. 164184.
 PR 10-JAN-1996: JP-201120.
 PR 15-JUL-1994: JP-136053.
 PR 17-JUN-1994: JP-136053.
 PA (SUNO) SUMITOMO CHEM CO LTD.
 DR WPI: 96-18231/19.
 DR PUBMED 814651.
 DR Antisense RNA inhibiting human derived cytochrome P450C28 - allows
 PT specific detection of cytochrome P450 species in humans
 PT Example 1; Pages 10-12; 13pp; Japanese.
 CC The present sequence encodes the human derived cytochrome (HDC)
 CC cytochrome P450C28. The gene encodes a protein of 290 amino
 CC acids. The protein is a member of the cytochrome P450 family.
 CC were transfected with an expression vector. The resulting HDC
 CC cultured and then disrupted to give a microsomal fraction. The
 CC HDC was purified from the fraction, and used to immunise and
 CC sensitise a mammal. Blood was drawn from the mammal, and an
 CC antiserum was prepared. The antiserum was used to detect the
 CC P450C28, partic. at serum dilution rates of 1:10000, and is
 CC substantially without cross reaction to other HDC P450 spp.
 SQ Sequence 1473 BP; 425 A; 335 C; 313 G; 400 T;
 Query Match 1 39; Score 15; DB 22; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951

CC (either as a fused protein or as a cell extract) and analysing the
 CC the amino acid sequence of the protein. The amino acid sequence
 CC or not rendered carcinogenic or "unsafe" if it is not detoxified or is
 CC metabolised to a carcinogenic cpd.
 SQ Sequence 1473 BP; 424 A; 335 C; 313 G; 401 T;
 Query Match 1 39; Score 15; DB 22; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951
 RESULT 237
 ID T17394 standard; cDNA; 1473 BP.
 AC T17394;
 DT 15-NOV-1995 (first entry)
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW Yeast WADPH-P450 reductase; safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism; ds.
 OS Homo sapiens.
 FT cd5 Location/Qualifiers
 FT 1..1473 /tag= a
 FT 108027198-A.
 PD 22-MAR-1995. 111298.
 PR 20-JUL-1993: JP-201120.
 PR 21-JUL-1993: JP-180246.
 PR 30-JUL-1993: JP-208279.
 PA (HAYASHI) HAYASHI K.
 PA (HAYASHI) HAYASHI K.
 PI Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T;
 Yabushiki Y;
 DR WPI: 95-116981/16.
 DR PUBMED 872370.
 DR Evaluation of safety of a chemical cpd. - using recombinant yeast
 PT expressing human cytochrome P450 and a yeast WADPH-P450 reductase
 PT Examples; Page 61-63; 124pp; English.
 CC The nucleotide sequence of the cDNA coding region for the human auxiliary
 CC cytochrome P450 species 2C8. The gene encodes a protein of 290 amino
 CC acids. The protein is a member of the cytochrome P450 family.
 CC product was cloned into the yeast expression vectors pAHEN or pAHEN to
 CC produce the vectors p2C8 for the expression of the cytochrome P450 alone
 CC or p2C8 for co-expression with the yeast WADPH-P450 reductase.
 CC The vectors are used in a method for evaluating the safety of a chemical
 CC cpd. The cpd is co-expressed with the yeast WADPH-P450 reductase and
 CC human cytochrome P450 molecular species 1A2 (087714), 2C9 (087715), 2B1
 CC (087716), or 3A4 (087717) or their auxiliary species and variants
 CC (087718-32). The yeast WADPH-P450 reductase, either as a fused protein or
 CC in cell extracts, and analysing the resulting metabolite to assess the
 CC whether the chemical compound, or its metabolite, will be converted into
 CC a carcinogenic or mutagenic form through metabolism in the liver.
 SQ Sequence 1473 BP; 425 A; 335 C; 313 G; 400 T;
 Query Match 1 39; Score 15; DB 15; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951
 RESULT 238
 ID T28330 standard; cDNA; 1473 BP.
 AC T28330;
 DT 15-NOV-1995 (first entry)
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 DE Human auxiliary cytochrome P450 species 2C8 coding region.
 KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW Yeast WADPH-P450 reductase; safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism; ds.
 OS Homo sapiens.
 FT cd5 Location/Qualifiers
 FT 1..1473 /tag= a
 FT 108027198-A.
 PD 22-MAR-1995. 111298.
 PR 20-JUL-1993: JP-201120.
 PR 21-JUL-1993: JP-180246.
 PR 30-JUL-1993: JP-208279.
 PA (HAYASHI) HAYASHI K.
 PA (HAYASHI) HAYASHI K.
 PI Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T;
 Yabushiki Y;
 DR WPI: 95-116981/16.
 DR PUBMED 872370.
 DR Evaluation of safety of a chemical cpd. - using recombinant yeast
 PT expressing human cytochrome P450 and a yeast WADPH-P450 reductase
 PT Examples; Page 70-72; 124pp; English.
 CC The nucleotide sequence of the cDNA coding region for the human auxiliary
 CC cytochrome P450 species 2C8. The gene encodes a protein of 290 amino
 CC acids. The protein is a member of the cytochrome P450 family.
 CC product was cloned into the yeast expression vectors pAHEN or pAHEN to
 CC produce the vectors p2C8 for the expression of the cytochrome P450 alone
 CC or p2C8 for co-expression with the yeast WADPH-P450 reductase.
 CC The vectors are used in a method for evaluating the safety of a chemical
 CC cpd. The cpd is co-expressed with the yeast WADPH-P450 reductase and
 CC human cytochrome P450 molecular species 1A2 (087714), 2C9 (087715), 2B1
 CC (087716), or 3A4 (087717) or their auxiliary species and variants
 CC (087718-32). The yeast WADPH-P450 reductase, either as a fused protein or
 CC in cell extracts, and analysing the resulting metabolite to assess the
 CC whether the chemical compound, or its metabolite, will be converted into
 CC a carcinogenic or mutagenic form through metabolism in the liver.
 SQ Sequence 1473 BP; 425 A; 335 C; 313 G; 400 T;
 Query Match 1 39; Score 15; DB 22; Length 1473;
 Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1123 agtttcgaactac 1137
 QY 937 AGTTTCGAACCTAC 951

Page 248

RESULT 242
ID TS9845 standard; DNA: 1512 BP.
AC TS9845.

```

Matches      15; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
Best Local Similarity 100.0%; Pred. NO. 8.37e+01;

```


CC from a phase lambda gt10 mouse embryo cDNA library. DNA of sequence
CC G5699 isolated from a cDNA library. DNA of sequence
CC cDNA as probe, and included the transcriptional regulatory element
CC of tek comprising the initiation codon and untranslated sequences.
CC The element may be used in gene therapy to introduce foreign genes
CC into endothelial cells to correct/prevent vascular disease.
SQ Sequence 1599 BP; 461 A; 327 C; 391 G; 411 T;

Query Match 1.38; Score 14; DB 10; Length 1599;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 216 aggtattttacc 229
|||||
QY 108 AGGCTATTGTACC 121

RESULT 253

ID 777094 standard; cDNA to mRNA; 1593 BP.
AC 777094-1998 (first entry)
DE Human N-acetylglucosaminyl transferase III gene.
KW N-acetylglucosaminyl transferase III; GAT-III; rat; human;
KW virus; replication; inhibitor; hepatitis B; hepatitis C; HIV;
KW viral disease; human immunodeficiency virus; da.
PB Rev aptens. Location/Qualifiers
FT CDS 1..1593
FT /tag= a
FT /note= "no stop codon given"
PN M09718816-1.
PF 29-MAY-1997.
PR 17-JUL-1996; J01886.
PS 17-NOV-1995; JP-322474.
PI Chiba Y, Miyoshi E, Taniguchi N;
DR WPI: 97-297877/27.
DR P-PSDB: 824015.
PT Virus replication inhibitor containing N-acetylglucosaminyl
PT transferase gene. Useful in treatment of hepatitis B, C
PT and HIV viral infections.
PS Claim 6; Page 22-26; 12pp; Japanese.
CC The present sequence represents human N-acetylglucosaminyl transferase
CC III (GAT-III) gene. This gene encodes a protein which is a
CC replication inhibitor of hepatitis B virus and hepatitis C virus.
CC The virus replication inhibitor is used against hepatitis B, C and HIV.
CC It is used for the treatment of viral diseases.
SQ Sequence 1593 BP; 269 A; 571 C; 488 G; 265 T;
Query Match 1.38; Score 15; DB 35; Length 1593;

Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 395 aggaagctctagg 409
|||||
CP 1015 AGGAGAGCTCTAGG 1001

RESULT 254

ID 729199 standard; cDNA to mRNA; 1593 BP.
AC 729199-1996 (first entry)
DE Rat N-acetylglucosaminyl transferase-III gene. Inhibitor of metastasis.
KW N-acetylglucosaminyl transferase-III; metastasis inhibitor; cancer;
KW neoplasia; invasive; da.
PB Ratius sp. Location/Qualifiers
FT CDS 1..1593
FT /tag= a
FT /note= "no stop codon"
FT 708109130-A.
PF 30-APR-1996.
PR 12-OCT-1994; 271802.
PS (TAKI) TAKARA SEIKO CO LTD.
DR P-PSDB: 824015.
PT Inhibitor of cancer metastasis - contains N-acetylglucosaminyl
PT transferase-III isolated from rat kidneys
PS Claim 3; Page 8-10; 12pp; Japanese.
CC The present sequence represents rat N-acetylglucosaminyl
CC transferase-III (GAT-III) gene. This gene encodes a protein which is
CC are useful to reinforce the activity of GAT-III present in cancer
CC cells and cells surrounding cancer cells for the inhibition of
CC metastasis, preventing the spread of cancer cells to tissues other
CC than the site of their origin. The clones were isolated from a rat
CC kidney homogenate.
SQ Sequence 1593 BP; 269 A; 568 C; 491 G; 265 T;
Query Match 1.38; Score 15; DB 23; Length 1593;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 395 aggaagctctagg 409
|||||
CP 1015 AGGAGAGCTCTAGG 1001

RESULT 255

ID 850731 standard; cDNA; 1599 BP.
AC 850731-1997
AC 850733;

DE 18-JUN-1990 (first entry)
DE Genomic DNA from Plasmodium falciparum K1 isolates encoding the 41 MD
DE merozoite antigen
KW Antimalaria vaccine; merozoite antigen; epitope; immunogen;
KW Antimalaria vaccine; passive immunisation; malaria diagnosis.
FH Plasmodium falciparum.
FS Key Location/Qualifiers
FT cds 346..1434
FT conflict 619..622
FT /tag= b
FT /note= "Differs from cDNA from M25 isolate of
FT P. falciparum."
FT /tag= c
FT /note= "Differs from cDNA from M25 isolate of
FT P. falciparum"
FT conflict 949..952
FT /tag= d
FT /note= "Differs from cDNA from M25 isolate of
FT P. falciparum"
PN PP-309746-A.
PR 27-SEP-1988; 114015.
PS 08-SEP-1987; CR-003486.
PA (Hoff) Hoffmann-La Roche AG.
PI Certi U; 101095/14.
DR P-PSDB: 893566.
PT New peptide(s).
PT contg. epitope(s) of Plasmodium falciparum merozoite antigen,
PT for use in malaria vaccines
CC The present sequence represents the cDNA of the 41 MD merozoite
CC antigen of Plasmodium falciparum K1. This cDNA is largely identical to this
CC genomic isolate of K1, whereas cDNA from M25 isolates differs
CC in 3 codons (see n82937). Plasmodium falciparum merozoite antigen can be
CC coupled to an affinity peptide, or adsorbed or covalently
CC to a solid support. The peptide, or adsorbed or covalently
CC to a solid support, is useful as immunogens in
CC anti-malaria vaccine. Antibodies directed against it are useful for
CC passive immunisation and diagnosis.
SQ Sequence 1599 BP; 619 A; 207 C; 259 G; 514 T;

Query Match 1.38; Score 14; DB 1; Length 1599;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1144 ggtctgtatttt 1157
|||||
CP 546 GGTCTGTATTITTT 533

RESULT 256
ID 062130 standard; cDNA; 1601 BP.
AC 062130-1994 (first entry)
DE 2A-MOV-1994 (first entry)
DE Endothelial specific receptor tyrosine kinase (tek) gene; mouse;
KW Endothelial specific receptor tyrosine kinase; tek gene; mouse;
KW murine; chromosome 4; cardiogenesis; as.
OS Mus musculus.
PB Key Location/Qualifiers
FT CDS 1..1601
FT /tag= a
FT /note= "open reading frame - does not begin with
FT initiation codon. Thought to encode the
FT cytoplasmic portion of a transmembrane RTK
FT consisting of a cytoplasmic domain followed by a short C-terminal tail (33 amino acids)."
FT 3'utr 908..1601
FT /tag= b
FT 032085291-A.
PF 31-JAN-1994.
PR 14-DEC-1992; 085291.
PS 10-JUL-1992; US-521795.
PA (MOON) MOON SINA HOSPITAL CORP.
DR 94-126598/14.
DR P-PSDB: 853147.
PT Receptor tyrosine kinase isolated from cells of the endothelial
PT lineage - used for identifying (ant)agonists of ligand-receptor
PT interactions.
PS Example 1; Fig 2; 51pp; English.
CC 062130 shows the nucleotide sequence of a 1.6kb tek cDNA isolated
CC from a 13.5 day mouse embryo cDNA library. Translation of this
CC sequence revealed a single large ORF that terminates with a TAG
CC stop codon. The deduced amino acid sequence suggests that the 684
CC features of the deduced amino acid sequence suggest that the 684
CC tek cDNA encodes the cytoplasmic domain of a transmembrane RTK,
CC consisting of the catalytic domain followed by a C-terminal tail.
SQ Sequence 1601 BP; 463 A; 335 C; 392 G; 411 T;

Query Match 1.38; Score 14; DB 11; Length 1601;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 216 aggtattttacc 229
|||||
QY 108 AGGCTATTGTACC 121

RESULT 257
ID V04252 standard; cDNA; 1604 BP.
AC V04252;

```

DT 05-JUN-1998 (first entry)
KW wild tomato acyltransferase cDNA
OS glucose acyl ester; pest repellent; emulsifier; emollient; ds.
FH Key Location/Qualifiers
FT 1..1395
FT /product= acyltransferase
FT
FT WO9748911-A1.
FT 24-DEC-1997. H11005
FT 21-JUN-1996; US-655966
FT (CORR.) CORNELL RES FOUND INC.
FT Changas GS, Steffens JC;
FT WPI: 98-06119/06.
FT
FT New acyltransferase from a wild tomato species - used to produce
FT glucose palmitoyl esters for use as pest repellents, emulsifiers and
FT emollients
FT Claim 3; Pages 10-24; 34pp; English. tomato acyltransferase, which
FT can be used to produce glucose palmitoyl esters, which
FT can be used to produce glucose palmitoyl esters, which
FT 1-O-palmitoyl-beta-D-glucose with glucose or its partial palmitoyl
FT esters. Disclosed is a similar process for glucose acyl esters in
FT general. Keyated sugars are useful as pest repellents in
FT cosmetics, emulsifiers in foods or cosmetics and emollients in
FT cosmetics.
FT
FT Sequence 1604 BP; 515 A; 298 C; 325 G; 466 T;
FT
FT Query Match 1.3%; Score 15; DB 39; Length 1604;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 775 gatcttcgaaga 789
CP 963 CACTTCCGACGACG 949

```

```

FT sig_peptide 91..198
FT /tag= b
FT mat_peptide 199..1524
FT /product= Bat-PA(H)
FT /note= Claim 12.
FT
FT EP-352119-A.
FT 24-JUN-1990. 107411
FT 20-JUN-1988; US-377221.
FT (MERRI) Merck and Co.
FT Duong LT, Jacobs JW, Friedman PA, Dixon RAP, Gardell BJ, Mark GE,
FT WPI: 98-02461/04.
FT P-PSDB: R05122/23.25.
FT Vampire bat glycosylated plasminogen activating protein - which needs
FT fibrin cofactor to activate plasminogen, has greater selectivity for
FT human plasminogen than PA.
FT Discloure: Fig 9; 18pp; English.
FT The sequence is that of the longest cDNA clone isolated from a library
FT of vampire bats. An additional 800 nt of 3' UT sequence ending
FT in a poly(A) tail is also included. The cDNA encodes a 260 amino
FT acid protein which, unlike tPA, does not contain the Kringle 2 domain and
FT plasmin-sensitive processing site. Three distinct species of activator
FT have been isolated corresponding to full-length (tag), finger-minus,
FT and finger-expressing forms of tPA. These are designated Bat-PA(H),
FT Bat-PA(F), and Bat-PA(L). The Bat-PA(H) is the most active. The new
FT proteins have greater selectivity towards fibrin-bound PA.
FT
FT Sequence 1620 BP; 434 A; 410 C; 440 G; 336 T;
FT
FT Query Match 1.3%; Score 15; DB 1; Length 1620;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 649 tcaaacctctggc 663
CP 686 TCAAAACCTCTGGTC 700

```

```

FT Key Location/Qualifiers
FT cds 1..1638
FT /tag= a
FT /product= ced-4_gene_product
FT
FT WO9320237-A.
FT 14-OCT-1993.
FT 01-APR-1993; D03102.
FT 01-APR-1992; US-861458.
FT (CORR.) GENENTECH INC.
FT Johnson CD, Marchionni MA,
FT WPI: 93-33693/42.
FT
FT P-PSDB: R42742.
FT
FT Long distance homology cloning of genes from lower organisms -
FT amino acid sequences
FT
FT Discloure: Fig 9; 18pp; English.
FT The primers/probes (Q49266-Q49295) are used to isolate the ced-4
FT gene from the nematode C. briggsae.
FT
FT Sequence 1638 BP; 482 A; 357 C; 343 G; 456 T;
FT
FT Query Match 1.3%; Score 14; DB 9; Length 1638;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 806 tcatcaccaccaga 819
CP 181 TCATCACCACCAGA 188

```

```

FT Key Location/Qualifiers
FT cds 141..1208
FT /tag= a
FT
FT G0270340-A.
FT 04-JAN-1995.
FT 15-JUN-1994; 012000.
FT 25-JUN-1993; US-083944.
FT (MERRI) MERCK & CO INC.
FT WPI: 95-024905/04.
FT
FT P-PSDB: R64226.
FT
FT New human 38 KD FK-506 binding protein and DNA encoding it -
FT useful for identifying cpds, able to block allo:graft rejection
FT in human.
FT Claim 2; Pages 33-34; 44pp; English.
FT Q80319 encodes R64226 the human 38kDa FK-506 binding protein. The
FT protein can be used to identify cpds, with immunosuppressive
FT activity, able to block allograft rejection by preventing T-cell
FT activation.
FT Identify, quantify and purify FK-506 and FK-506 like cpds
FT
FT Sequence 1641 BP; 324 A; 550 C; 507 G; 260 T;
FT
FT Query Match 1.3%; Score 14; DB 13; Length 1641;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 828 cacttcggcggc 841
CP 452 CACTTCGGCGGCGC 465

```

```

FT Key Location/Qualifiers
FT cds 1..1638
FT /tag= a
FT /product= ced-4_gene_product
FT
FT WO9320237-A.
FT 14-OCT-1993.
FT 01-APR-1993; D03102.
FT 01-APR-1992; US-861458.
FT (CORR.) GENENTECH INC.
FT Johnson CD, Marchionni MA,
FT WPI: 93-33693/42.
FT
FT P-PSDB: R42742.
FT
FT Long distance homology cloning of genes from lower organisms -
FT amino acid sequences
FT
FT Discloure: Fig 9; 18pp; English.
FT The primers/probes (Q49266-Q49295) are used to isolate the ced-4
FT gene from the nematode C. briggsae.
FT
FT Sequence 1638 BP; 482 A; 357 C; 343 G; 456 T;
FT
FT Query Match 1.3%; Score 14; DB 9; Length 1638;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 806 tcatcaccaccaga 819
CP 181 TCATCACCACCAGA 188

```

```

FT Key Location/Qualifiers
FT cds 141..1208
FT /tag= a
FT
FT G0270340-A.
FT 04-JAN-1995.
FT 15-JUN-1994; 012000.
FT 25-JUN-1993; US-083944.
FT (MERRI) MERCK & CO INC.
FT WPI: 95-024905/04.
FT
FT P-PSDB: R64226.
FT
FT New human 38 KD FK-506 binding protein and DNA encoding it -
FT useful for identifying cpds, able to block allo:graft rejection
FT in human.
FT Claim 2; Pages 33-34; 44pp; English.
FT Q80319 encodes R64226 the human 38kDa FK-506 binding protein. The
FT protein can be used to identify cpds, with immunosuppressive
FT activity, able to block allograft rejection by preventing T-cell
FT activation.
FT Identify, quantify and purify FK-506 and FK-506 like cpds
FT
FT Sequence 1641 BP; 324 A; 550 C; 507 G; 260 T;
FT
FT Query Match 1.3%; Score 14; DB 13; Length 1641;
FT Best Local Similarity 100.0%; Pred No. 347e+02;
FT Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
DB 828 cacttcggcggc 841
CP 452 CACTTCGGCGGCGC 465

```

```

JC virus, antigenprotein, VP23, VP2311, VP3, VP1, contour protein,
progressive multifocal leukoencephalopathy, aa.
KW 00224658-A.
OS 09
PN 23-FEB-1986; JP-062167.
PR 23-FEB-1986; JP-062167.
PR 23-FEB-1986; JP-062167.
PR (AISE) A IASIN SEIKI KK.
PR WT: 97-48409/43.
PR 23-FEB-1986; JP-062167.
PR (AISE) A IASIN SEIKI KK.
PR used in diagnosis of progressive multifocal leukoencephalopathy
PR class 3; Fig 5-7; 15pp; Japanese.
A method has been developed for the preparation of JC virus particles -
containing the genome of JC virus particles, VP23, VP2311, VP3,
VP1, VP11, VP12, and VP13. The method involves the preparation of a
region between bases 277 and 493 from the replication starting point is
deleted to prepare a virus particle-producing DNA fragment. Then the
virus particle-producing DNA fragment is recombinant to an expression
vector to produce JC virus particles. The present sequence represents a JC virus
DNA fragment of the present invention. The method can prepare JC virus
particles or JC virus contour protein in a high yield in a short period.
Sequence 1630 BP: 513 A; 298 C; 366 G; 465 T;
Query Match 1.3%; Score 15; DB 38; Length 1650;
Best Local Similarity 100.0%; Pred. Nos. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 980 gtscgaggaagccgc 994
Cp 897 GTGCAGAAAGCCG 883

RESULT 264
ID T16092 standard; cDNA to mRNA; 1650 BP.
C 16092.1986 (first entry)
D2 KM1-7 precursor cDNA clone 11-7.
KW KM1-7, glutathione reducing protein; nuclear inclusion a;
KW protease; autolysis; protein fusion; cleavage; chloroindophenol;
KW 5S rRNA; stress; activated oxygen; therapy; ds.
OS Homo sapiens.
PR Key Location/Qualifiers
FF signal_peptide 1..69
FF tag_a 74tag_a
FF tag_b 74tag_b
FF mat_peptide 74tag_a
FF A0954970-A.
PR 25-JUN-1986. 034970
PR 13-DEC-1984; JP-161053

```

```

DNU m3). encoding human nuclear factors of activated T cells -
useful for screening potential therapeutic and diagnostic agents for
immune system diseases
PT Diaclosure; Page 55-56; 6pp; English.
PS A cDNA sequence (J36871) comprises the 3' end of human nuclear
CC factor NFATc1 (NFATc1), also known as NFAT4, which has been shown
CC to activate transcription of NFAT target genes in lymphocytes and
CC alpha 700-1068 (M02523). 2 Other classes of NFAT4 cDNA, NFAT4a
CC (J33677) and NFAT4b (J36975), were also isolated; the NFAT4c
CC transcript was most abundant. The major sites for NFAT4 expression
CC were identified in muscle and thymus. NFAT cDNAs (see also J36865-68)
CC are related to other members of the NFAT family of proteins, which
CC human peripheral blood lymphocytes by screening with a probe derived
CC from the rel similarity region of the murine NFATp gene. NFAT genes
CC can be used for prodn. of recombinant NFATs, as probes and primers,
CC and for studying the regulation of NFAT gene expression and intracellular
CC availability of active NFAT. 401 A: 418 C: 354 G: 489 T:
SQ Query Match 1330 Score 14; DB 29; Length 1662;
DB Best Local Similarity 100.0%; Pos No 374+02;
Matches 14; Conservative O; Mismatches 0; Indels 0; Gaps 0;
Db 1447 atttgcttttata 1460
Qv 183 CTTCGCTTGTATA 196

RESUUT_266
IC T36871 standard; cDNA, 1662 BP.
ID T36871;
DT DE NOV-1996 (first entry)
DT 18-NOV-1996 (last entry)
KW Human transcription factor NFAT4c (3' end) cDNA.
KW transcription factor NFAT4c; NFAT; NFAT4c;
KW transcription factor-binding domain; Immune system disease;
KW gene therapy; diagnosis; ds.
OS Homo sapiens.
SS Key Feature Location/Qualifiers
FT Key Feature /tag=
FT Cds /product= NFAT4a
FT F1
FT F2
FT F3 W0526959-AI.
FT F4 CDNA
FT F5 CDNA
FT F6 CDNA MAR-1996; W03113
FT F7 CDNA MAR-1996; US-395679.
FT F8 CDNA MAR-1995; US-395679.
FT F9 (TUL-) TULARIK INC.
FT F10 Hoey 56-412738/A1.
FT F11 PI
FT F12 PI
FT F13 P-FSBI; M02253.
FT F14
FT F15 DNU m3). encoding human nuclear factors of activated T cells -
FT useful for screening potential therapeutic and diagnostic agents for
```

Disclosures: Page 55-56, 64pp, English
 A cDNA sequence (T16871) comprises the 3' end of human nuclear
 factor of activated T-cells class 4c (NFAT4c) cDNA, coding for amino
 acids 700-1068 (W02323). 2 Other classes of NFAT4 cDNAs, NFAT4a
 (T13677) and NFAT4b (T16873), were also isolated: the NFAT4c
 cDNA encodes for 368 amino acids (MW 42,666), the NFAT4a cDNA
 is in skeletal muscle and thymus, NFAT4b cDNA for NFATc2 (565-649)
 were isolated from cDNA libraries prep'd from Jurkat T-cells and
 human peripheral blood lymphocytes by screening with a probe derived
 from the similarity region of the murine NFATc gene. NFATc genes
 from Jurkat T-cells and peripheral blood lymphocytes, as probes and primer,
 and (esp. antisense) to modulate cellular expression of intracellular
 availability of active NFAT.
 Sequence 1662 BP 401 A: 418 C: 354 G: 489 T:

```

Query Match          1.3%  Score 14;  DB 23;  Length 1662;
Best Local Similarity 100.0%;  Pred. No. 3,47+02;
Matches 14;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db  911 ttccactgatggg 924
Cp  605 TTCACCTGATGGG 592

```

	RESULT	267
ID	T85666 standard; DNA; 1667 bp.	
AC	785666_1997 (4 gene survey)	
DE	CPT-SPSS C4' gene fusion	
KW	expression cassette; inducible promoter; alcA; alcB; regulator;	
KW	N-phosphonomethyl-L-glutamic acid; phosphonate; C4'	
KW	5-enol-pyruvylshikimate 3-phosphate C4'; chloroplast transit sequence ;	
OS	Chimeric Arabidopsis CPT1; dicotyledon; monocotyledon; crop; ss.	
OS	Chimeric Petaluma hybrida.	
PP	705P521894	
PP	705P521894	

PP 02-AUG-1996: G01883.
PP 03-AUG-1995: GB-01941.
PA (ZERN) ZENDECA LTD.
DA WPI: 97-154273/14.
PP Chemically inducible cassette for expressing herbicide resistance
PP gene in plants - and derived plants, partic. for resistance to
PP glyphosate avoids constitutive expression and minimises development
PP of resistance to glyphosate
PP Example 3: Fig 10: 59pp; English.
PS

CC A chemically inducible plant gene expression cassette which comprises an
CC inducible promoter linked to a gene (I) that confers resistance to a
CC herbicide, is claimed. (I) imparts resistance to N-phosphonomethyl-

Ft		preferred for use in constructs of the
Pt		invention (Claim 5).
Pt	primer_bind	complement (\$22..\$40)
Pt		/note= "Primer HP20H binding site"
Pt		/date= 1667
Pt	primer_bind	/cseq= f
Pt		/note= "Primer HP20C binding site"
Pt		W09138475-A1.
Pt	PD	05-DEC-1996.
Pt	PD	80797
Pt	PD	OJ-JUN-1997; S2007007
Pt	PD	OJ-JUN-1997; S2007007
Pt	PD	21-MAR-1996; S96-001085.
Pt	PA	(ASTR) ASTRA AB.
Pt	PA	(ENGLAND) ENGLAND A;
Pt	WI	WIPI_97-034307/OI
Pt	DOR	

P-PDDB: 807449.
Helicobacter pylori surface exposed antigen - useful for treatment,
diagnosis and vaccine development in gastric infections
Claim 5, Page 32-34; 49pp; English
A cDNA clone (743834) codes for a 39 kDa surface-exposed antigen
(807449) of the helicobacter pylori that is useful for the treatment
of gastric diseases and for the development of vaccines and
vaccines. Genomic fragments of H. pylori were cloned into a
CC Express vector. Following in vitro packaging, the library was
CC screened by infecting strain A1.1 Blue Mx2, and plated onto IPG
CC Y-agar. Positive clones were identified by immunoblotting
CC detected using monoclonal antibody H930-1.1. 2 Clones, previously
CC 743835, differing only at position 1438, were identified. Antigen
CC nucleic acids can be used to produce recombinant antigen in host
CC bacteria. 1570 bp; 1997 N. 259 C. 371 G. 443 F.
CC

```

Query Match      1.3%   Score 15; DB 25; Length 1670;
Best Local Similarity 100.0%   Pred. No. 8,376+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      250 ggataaagaaccca 264
CP      650 GGCTTAAGAAACCA 636

```

[illegible]

glycine (glyphosate) or its salts, but may also be a gene for resistance to other herbicides. The inducible promoter (e.g. aca, acah, or other aca-induced gene promoter) is operatively linked to an acaR regulator expression cassette. Induced expression of (1) avoids the risk that constitutive expression of acaR will confer resistance to acaR herbicides, and (2) is controlled by herbicide applied without inducer and minimizes the chance that herbicide-resistant strains of weeds will arise. The expression cassette is strictly controlled and suitable for general use in plants with genes conferring resistance to glyphosate. The switch was inserted into inducible expression of glyphosate oxidase (GOX) in plants. GOX switchable GOX was expressed alone or in conjunction with constitutive expression of acaR from a non-tyrosine kinomate[®] phosphate (PS899) CMA. Constructs were ordered and synthesized by the University of California, San Diego. The present sequence comprises a fusion of the BPSPS class 1 gene from *Petunia* hybrid fused to the chloroplast transit sequence 2 gene from *Arabidopsis thaliana* (CPRI). This sequence was ligated into pMB31 (see Example 1567 BP; pMB31: 4897-4903, 4903-4910, 4910-4920, 4920-4929, 4929-4936, 4936-4941, 4941-4947, 4947-4953, 4953-4959, 4959-4965, 4965-4971, 4971-4977, 4977-4983, 4983-4989, 4989-4995, 4995-5001, 5001-5007, 5007-5013, 5013-5019, 5019-5025, 5025-5031, 5031-5037, 5037-5043, 5043-5049, 5049-5055, 5055-5061, 5061-5067, 5067-5073, 5073-5079, 5079-5085, 5085-5091, 5091-5097, 5097-5103, 5103-5109, 5109-5115, 5115-5121, 5121-5127, 5127-5133, 5133-5139, 5139-5145, 5145-5151, 5151-5157, 5157-5163, 5163-5169, 5169-5175, 5175-5181, 5181-5187, 5187-5193, 5193-5199, 5199-5205, 5205-5211, 5211-5217, 5217-5223, 5223-5229, 5229-5235, 5235-5241, 5241-5247, 5247-5253, 5253-5259, 5259-5265, 5265-5271, 5271-5277, 5277-5283, 5283-5289, 5289-5295, 5295-5301, 5301-5307, 5307-5313, 5313-5319, 5319-5325, 5325-5331, 5331-5337, 5337-5343, 5343-5349, 5349-5355, 5355-5361, 5361-5367, 5367-5373, 5373-5379, 5379-5385, 5385-5391, 5391-5397, 5397-5403, 5403-5409, 5409-5415, 5415-5421, 5421-5427, 5427-5433, 5433-5439, 5439-5445, 5445-5451, 5451-5457, 5457-5463, 5463-5469, 5469-5475, 5475-5481, 5481-5487, 5487-5493, 5493-5499, 5499-5505, 5505-5511, 5511-5517, 5517-5523, 5523-5529, 5529-5535, 5535-5541, 5541-5547, 5547-5553, 5553-5559, 5559-5565, 5565-5571, 5571-5577, 5577-5583, 5583-5589, 5589-5595, 5595-5601, 5601-5607, 5607-5613, 5613-5619, 5619-5625, 5625-5631, 5631-5637, 5637-5643, 5643-5649, 5649-5655, 5655-5661, 5661-5667, 5667-5673, 5673-5679, 5679-5685, 5685-5691, 5691-5697, 5697-5703, 5703-5709, 5709-5715, 5715-5721, 5721-5727, 5727-5733, 5733-5739, 5739-5745, 5745-5751, 5751-5757, 5757-5763, 5763-5769, 5769-5775, 5775-5781, 5781-5787, 5787-5793, 5793-5799, 5799-5805, 5805-5811, 5811-5817, 5817-5823, 5823-5829, 5829-5835, 5835-5841, 5841-5847, 5847-5853, 5853-5859, 5859-5865, 5865-5871, 5871-5877, 5877-5883, 5883-5889, 5889-5895, 5895-5901, 5901-5907, 5907-5913, 5913-5919, 5919-5925, 5925-5931, 5931-5937, 5937-5943, 5943-5949, 5949-5955, 5955-5961, 5961-5967, 5967-5973, 5973-5979, 5979-5985, 5985-5991, 5991-5997, 5997-6003, 6003-6009, 6009-6015, 6015-6021, 6021-6027, 6027-6033, 6033-6039, 6039-6045, 6045-6051, 6051-6057, 6057-6063, 6063-6069, 6069-6075, 6075-6081, 6081-6087, 6087-6093, 6093-6099, 6099-6105, 6105-6111, 6111-6117, 6117-6123, 6123-6129, 6129-6135, 6135-6141, 6141-6147, 6147-6153, 6153-6159, 6159-6165, 6165-6171, 6171-6177, 6177-6183, 6183-6189, 6189-6195, 6195-6201, 6201-6207, 6207-6213, 6213-6219, 6219-6225, 6225-6231, 6231-6237, 6237-6243, 6243-6249, 6249-6255, 6255-6261, 6261-6267, 6267-6273, 6273-6279, 6279-6285, 6285-6291, 6291-6297, 6297-6303, 6303-6309, 6309-6315, 6315-6321, 6321-6327, 6327-6333, 6333-6339, 6339-6345, 6345-6351, 6351-6357, 6357-6363, 6363-6369, 6369-6375, 6375-6381, 6381-6387, 6387-6393, 6393-6399, 6399-6405, 6405-6411, 6411-6417, 6417-6423, 6423-6429, 6429-6435, 6435-6441, 6441-6447, 6447-6453, 6453-6459, 6459-6465, 6465-6471, 6471-6477, 6477-6483, 6483-6489, 6489-6495, 6495-6501, 6501-6507, 6507-6513, 6513-6519, 6519-6525, 6525-6531, 6531-6537, 6537-6543, 6543-6549, 6549-6555, 6555-6561, 6561-6567, 6567-6573, 6573-6579, 6579-6585, 6585-6591, 6591-6597, 6597-6603, 6603-6609, 6609-6615, 6615-6621, 6621-6627, 6627-6633, 6633-6639, 6639-6645, 6645-6651, 6651-6657, 6657-6663, 6663-6669, 6669-6675, 6675-6681, 6681-6687, 6687-6693, 6693-6699, 6699-6705, 6705-6711, 6711-6717, 6717-6723, 6723-6729, 6729-6735, 6735-6741, 6741-6747, 6747-6753, 6753-6759, 6759-6765, 6765-6771, 6771-6777, 6777-6783, 6783-6789, 6789-6795, 6795-6801, 6801-6807, 6807-6813, 6813-6819, 6819-6825, 6825-6831, 6831-6837, 6837-6843, 6843-6849, 6849-6855, 6855-6861, 6861-6867, 6867-6873, 6873-6879, 6879-6885, 6885-6891, 6891-6897, 6897-6903, 6903-6909, 6909-6915, 6915-6921, 6921-6927, 6927-6933, 6933-6939, 6939-6945, 6945-6951, 6951-6957, 6957-6963, 6963-6969, 6969-6975,

```
DB      330 ccaaggctctcca 343
          ||||| |||||
CP      978 ccacagctcttcca 965

Query Match      1.3%   Score 14; DB 33; Length 1667;
Best Local Similarity 100.0%; Pred. No. 3,47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 769 standard: cDNA, 1670 BP.
770
771 AC 743834:
772 40-MAR-1997 (first entry)
773 Helicobacter pylori surface-exposed 33 kDa antigen cDNA.
774 DE antigen: chlamy; immunogen: vaccine; genetic immunisation: ulcer;
775 Helicobacter pylori strain CGU 17874. HPOC 11637.
776
777 AC 743834:
778 Helicobacter pylori strain CGU 17874. HPOC 11637.

FF	Key	Location/Qualifiers
FF	rbs	782..785
FF	cds	793..1575
FF		/tag= b
FF	note	(note) the nucleic acid represented by bases
FF	note	(note) the protein sequence
FF	note	(note) the mature protein coding sequence is
FF	signal_peptide	793..873
FF		/tag= c
FF	mat_peptide	874..1575
FF		/tag= d
FF		(note) the mature protein coding sequence is

```

rbs      782..785
          ATGTTTCTGGTAA
cds       793..1575
          /tag= b
/note= "the nucleic acid represented by bases
        793-1575 are part of the mature protein
        constructs of the invention (claim 5)"
signal_peptide 793..873
           /tags= c
           /start= 1
mat_peptide 873..1575
           /tag= d
/note= "the mature protein coding sequence is
        preferred for use in constructs of the
        invention (Claim 5)"
primer_bind 1327..1540
           /tag= e
           /note= "primer HP30N binding site"
primer_bind 1469..1487
           /tag= f
           /note= "primer HP30C binding site"

W09S18475-A1.
03-DIC-1996. SEQID77
01-JUN-1995; SE-002007.
21-MAR-1996; SE-001085.
(ASTR ) ASTRA AB.
PCT/SE95/00131. Stockholm A.
EP1 97 034307.03

```

P-F8B2; 907450.
Helicobacter pylori surface exposed antigen - useful for treatment,
diagnosis and vaccine development in gastric and duodenal infections
Claim 5, Page 37-38, 49pp; English, German, French, Japanese, Spanish
A CDNA clones (743835) coding for a 39 kDa surface-exposed antigen
907450) of Helicobacter pylori that is useful for the treatment
of gastric and duodenal infections. The antigen is a surface exposed
vaccine. Genomic fragments of H. pylori were cloned into a pAP
Express vector. Following in vitro packaging, the library was
titrated by infecting strain XL-1 blue MM, and plated onto IPKC
medium. The library was screened for the presence of the antigen using
detected using monoclonal antibody 907450-1.1.6. 2 clones were
743834), differing only at position 1450, were identified. Antigen
nucleic acid was sequenced and compared with the published sequence
of the antigen from H. pylori strain NCTC 11637. The antigen was
sequenced. 1970 BP; 388 AAs. 258 C; 371 G; 443 T.

```
Query Match      1.38; Score 15; DB 25; Length 1670;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches    15; Conservative    0; Mismatches 0; Indels 0; Gaps 0;
```


[illegible][illegible]

```

DE Plasmid pTnp-35 DNA encoding the 2,5-difetoglucanate-reductase gene.
KW 2,5-diketoglucanate-reductase; ascorbic acid, aa.
OS Corynebacterium.
FH Promoter
FT Location/Qualifiers
FT /feature_1=
FT /tag=a
FT /label=E. coli ttp promoter
FT /start=252..257
FT /stop=260
FT /tags=b
FT /tag=c
FT rbs
FT /start=279..283
FT /stop=283
FT /tags=c
FT cds
FT /start=296..1130
FT /stop=1130
FT /label=2,5-DKG-reductase gene
FT
PD EP-132108-A.
PN 10-JAN-1985.
PR 10-JUN-1983; US-508409
PR 28-JUN-1983; US-508428
PR 28-JUN-1983; US-508428
PR 28-JUN-1983; US-508410
PR 14-JUN-1984; US-620585
PR 14-JUN-1984; US-620585
PR 14-JUN-1984; US-620585
PR 14-JUN-1984; US-620585
PR (GENTECH) GENEITECH INC.
PA Esell DA, Light DR, Basteter WH, Lasarus BA, Miller JV;
PI P-5083, P-5094, P-5095.
PT New pure 2,5-di-keto-gluconic acid reductase - prepd. by
PT recombinant DNA methods for ascorbic acid prodn.
PT Disclousure, Fig 4; 44pp; English. Award by recombinant DNA techniques.
P5 2,5-DKG-reductase is used to convert 2,5-DKG stereoselectively.
CC 2-KLG, which is an intermediate in the production of ascorbic acid
CC (vitamin C).
CC Sequence 1853 BP; 327 A; 645 C; 587 G; 294 T;
SQ
Query Match 1:38; Score 14; DB 3; Length 1853;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1325 catcgcactgtctg 1338
|||||
OY 289 CATCGCTACTGTCTC 302

RESULT 269
ID N90611 standard; DNA; 1853 BP.
NC 290611.1980 (first entry)
DC 290611.1980 (first entry)
DE Sequence including 2,5-difetoglucanic acid reductase gene and vfrM15

```

DE Region of pTpi-35 expression vector contg. 2,5-DKG gene
 DF Anionic acid, 2-keto-L-gluconic acid, 2,5-diketogluconic acid; as.
 EY synthetic.
 FE Key
 FF Location/Qualifiers
 FT cda 297...1133
 FT cda 237...257
 FT promoter /tag= b
 FT rbs /label=coli trp promoter
 FT rbs /tag= c
 FT
 FN EP-29303-A.
 FP 23 NOV-1988. 084581
 FR 21 MAR-1987. 054530
 PR 21 MAR-1987. 054530
 PA (GTRR) Genentech Inc.
 PI Anderson S, Light DR, Marks C, Rastetter WH:
 PR WPJ: 88-31164/47.
 PT Prod'n. of 2-keto-gluconic acid - using microsome contg. transferred
 PT genetic material capable of conversion of glucose or other ordinary
 PT metabolic
 PT This was a p: English
 CC This was a p: English
 CC The enzymes' regions the 2,5-diketogluconic acid (DGT) reductase genes
 CC and the control regions of the pTpi-35 expression vector. The vector is
 CC used in the conversion of 2,5-DKG to 2-keto-L-gluconic acid which is an
 CC intermediate in the ascorbic acid biosynthesis pathway. The vector is
 CC used to transform E. coli cells which are deficient in the conversion of
 CC ordinary metabolites, eg glucose into the 2,5-DKG substrate for this
 CC enzyme. See also H92093 and H92100.
 SQ enzyme. 1854 BP: 327 A; 644 C; 588 G; 295 T;
 Query Match 1.3%; Score 14; DB 1; Length 1854;
 Match Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1326 catcgccatgctgc 1339
 | | | | | | | | | |
 | | | | | | | | | |
 QY 289 CATGCGCACTGCTGC 302
 RESULT 291
 ID T91737 standard; DNA: 1875 BP.
 AC T91737;
 DT 10/19/88 (first entry)
 DA 10 APR-1988
 DR 10 APR-1988
 JC virus; agnoprotein; VP3; VP21; VP3; VP4; contour protein;
 KW progressive multifocal leukoencephalopathy; as.
 CC JC virus.
 FT Location/Qualifiers
 FT CD8 138...875
 FT CD8 138...875

RESULT 295
ID Q23241 standard; cDNA: 1938 BP.
AC Q23241
DE Sequence encoding insulin-like growth factor binding protein
DE (IGFBP)-5.
OS Bone development; osteoporosis; anemia; growth; wound healing; ss.
FH Key Reptition. Location/Qualifiers
FT signal_peptide 24..86
FT mat_peptide 87..1800
FT /tag= a
FT /tag= b
FN M09203470-A.
PD 05-MAR-1992. U06140
PR 28-AUG-1991; US-574613.
PR 31-AUG-1990; US-577391.
PR (CHIR-) CHIRON CORP.
PI Kiefer MC, Maslars F, Zapf J, Born W.
DR WJ: 92-09681/12.
PT New insulin-like growth factor binding protein IGFBP-5 - for
PT diagnosis and treatment of cancer, osteoporosis, anemia and
PT abnormal growth
PT Disclosure; Fig 1: 65pp; English
CC The inventors claim a recombinant DNA mol. encoding IGF BP-5 or a
CC fragment, having greater than or equal to 10 nucleotides. The IGF
CC BP-5 is pref. of human origin and is genomic or cDNA pref. contained
CC in pBR322.5.
SQ Sequence 1938 BP: 369 A; 576 C; 584 G; 409 T;
Query Match 1.38; Score 14; DB 3; Length 1938;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 540 tctctccagagca 553
QY 832 tctctccagagca 845
RESULT 296
ID Q23274 standard; DNA: 1938 BP.
AC Q23274
DE Sequence encoding insulin-like growth factor binding protein 5
DE (IGFBP-5).
OS Bone development; growth; erythropoiesis; chondrogenesis; ss.
FH Key Reptition. Location/Qualifiers
FT signal_peptide 24..86
FT mat_peptide 87..1800
FT /tag= a
FT /tag= b
FN M09203471-A.
PD 05-MAR-1992. U06141
PR 28-AUG-1991; US-574613.
PR 31-AUG-1990; US-577391.
PR (CHIR-) CHIRON CORP.
PI Kiefer MC, Maslars F, Zapf J, Born W.
DR WJ: 92-09681/12.
PT New insulin-like growth factor binding protein IGFBP-5 - for
PT diagnosis and treatment of cancer, osteoporosis, anemia and
PT abnormal growth
PT Disclosure; Fig 1: 65pp; English
CC The inventors claim a recombinant DNA mol. encoding IGF BP-5 or a
CC fragment, having greater than or equal to 10 nucleotides. The IGF
CC BP-5 is pref. of human origin and is genomic or cDNA pref. contained
CC in pBR322.5.
SQ Sequence 1938 BP: 369 A; 576 C; 584 G; 409 T;
Query Match 1.38; Score 14; DB 3; Length 1938;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 540 tctctccagagca 553
QY 832 tctctccagagca 845

Key Location/Qualifiers
FH signal_peptide 24..86
FT mat_peptide 87..1800
FT /tag= a
FT /tag= b
FN M09203470-A.
PD 05-MAR-1992. U06140
PR 28-AUG-1991; US-574613.
PR 31-AUG-1990; US-577392.
PR (CHIR-) CHIRON CORP.
PI Kiefer MC, Maslars F.
DR WJ: 92-09681/12.
PT DNA encoding insulin-like growth factor binding protein-5 - used
PT to obtain large quantities of protein useful in e.g. wound
PT healing. Fig 1: 65pp; English
CC The inventors claim a recombinant DNA mol. encoding IGF BP-5 or a
CC fragment, having greater than or equal to 10 nucleotides. The IGF
CC BP-5 is pref. of human origin and is genomic or cDNA pref. contained
CC in pBR322.5.
SQ Sequence 1938 BP: 369 A; 576 C; 584 G; 410 T;
Query Match 1.38; Score 14; DB 3; Length 1938;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 540 tctctccagagca 553
QY 832 tctctccagagca 845
RESULT 297
ID T95211 standard; DNA: 1961 BP.
AC T95211
DE Human disintegrin metalloprotease (aggrucanase) gene.
DE Disintegrin; metalloprotease; arthritis; osteoarthritis; diagnosis;
KW aggrucanase; arthropathy; osteoporosis; ankylosing spondylitis;
GO post; Lyme disease; human; ss.
FH Key Reptition. Location/Qualifiers
FT CDS 2..1477
FT /tag= a
FN M0971931-A1.
PD 04-SEP-1997. U03217.
PR 28-FEB-1997; U03217.
PR 01-MAR-1996; US-012679.
PR (PROC) PROCTER & GAMBLE CO.
PA (UTCA-) UNIV CASE WESTERN RESERVE.

PI Hagui T, Tindal WH.
DR WPI: 97-448626/41.
DR P-PSDB; W35292.
DE New human disintegrin metalloprotease - is differentially expressed
DE in human arthritis.
PT Disintegrin metalloprotease-mediated diseases, e.g. osteoarthritis
PT Disclosure; Page 13-15; 29pp; English.
CC This isolated DNA fragment encodes a human disintegrin
CC metalloprotease (see W35292). Designated aggrucanase, that is
CC involved in the degradation of aggrecan, a major component of
CC and which is involved in tissue remodeling and breakdown. It was
CC isolated from primary cultures of interleukin-1 stimulated human
CC articular (femoral head) chondrocytes using differential display
CC analysis and was used to recover full length clones (see W35212).
CC The full length cDNA was subcloned into a pMT vector and
CC be used in the construction of recombinant expression vectors for
CC production of the metalloprotease in host cells. The gene provides
CC a marker to monitor the development, including the earliest
CC stages of the disease process, including osteoarthritis.
CC Inhibitors of the disintegrin including osteoarthritis-mediated
CC can potentially be used in the treatment of metalloprotease-mediated
CC diseases such as osteoarthritis, osteoporosis, ankylosing
CC spondylitis, rheumatoid or inflammatory arthritis, gout and Lyme
CC disease.
SQ Sequence 1961 BP: 608 A; 395 C; 432 G; 526 T;
Query Match 1.38; Score 14; DB 36; Length 1961;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1548 agtcatctctcca 1561
QY 498 agtcatctctcca 511
RESULT 298
ID T89344 standard; cDNA: 1977 BP.
AC T89344
DE Human p62 cDNA #2
DE Human p62 cDNA #2
KW p62; cytoplasmic; T cell; B cell; development; activation; modulation;
OS cellular response; cell proliferation; autoimmune disease; ss.
FH Key Reptition. Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= p62
FT /note= 'partial CDS'
FN M0972255-A1.
PD 26-JUN-1997.
PR 11-DEC-1996; U19944.

PR 19-DEC-1995; US-574959.
PA (DAND) DANA FARMER CANCER INST INC.
PI Jung I, Shin J, Strominger JL, Vladanudi RK.
DR WPI: 97-341351/31.
DR P-PSDB; W35292.
DE New human disintegrin metalloprotease - is differentially expressed
DE in human arthritis.
PT Disintegrin metalloprotease-mediated diseases, e.g. osteoarthritis
PT Disclosure; Page 13-15; 29pp; English.
CC This isolated DNA fragment encodes a human disintegrin
CC metalloprotease (see W35292). Designated aggrucanase, that is
CC involved in the degradation of aggrecan, a major component of
CC and which is involved in tissue remodeling and breakdown. It was
CC isolated from primary cultures of interleukin-1 stimulated human
CC articular (femoral head) chondrocytes using differential display
CC analysis and was used to recover full length clones (see W35212).
CC The full length cDNA was subcloned into a pMT vector and
CC be used in the construction of recombinant expression vectors for
CC production of the metalloprotease in host cells. The gene provides
CC a marker to monitor the development, including the earliest
CC stages of the disease process, including osteoarthritis.
CC Inhibitors of the disintegrin including osteoarthritis-mediated
CC can potentially be used in the treatment of metalloprotease-mediated
CC diseases such as osteoarthritis, osteoporosis, ankylosing
CC spondylitis, rheumatoid or inflammatory arthritis, gout and Lyme
CC disease.
SQ Sequence 1977 BP: 469 A; 531 C; 546 G; 431 T;
Query Match 1.38; Score 14; DB 36; Length 1977;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 940 agtcatctctcca 953
QY 510 agtcatctctcca 497
RESULT 299
ID M90225 standard; DNA: 1982 BP.
AC M90225
DE Malaria-specific Plasmodium falciparum protein.
DE Malaria-specific Plasmodium falciparum protein; vaccine.
OS Plasmodium falciparum; ss.
FH Key Reptition. Location/Qualifiers
FT CDS 526..1611
FT /tag= a
FN M0972255-A1.
PD 26-JUN-1997.
PR 11-DEC-1996; U19944.

CC	Open reading frame represented by 780102. This sequence, 780099, and 780100 all represent cDNA sequences of the invention. These sequences all have 100% identity to each other.	Score 14; E-Value 1.7e-04; Pos. No. 1-747; Length 740; Mismatches 0; Indels 0; Gaps 0;
CC	All K3-1 genes, and contain the open reading frames shown in 780102, 780104, and 780101, respectively. These genes, their products and their functions are described in the text.	Score 14; E-Value 1.7e-04; Pos. No. 1-747; Length 740; Mismatches 0; Indels 0; Gaps 0;
CC	Sequence of the proteinase of the Alzheimer's disease.	Score 626 A; 366 C; 424 G; 664 T;
CC	Sequence 2080 BP.	
CC	Query Match	Score 14; E-Value 1.7e-04; Pos. No. 1-747; Length 740; Mismatches 0; Indels 0; Gaps 0;
CC	Best Local Similarity 100.0%; Pos. No. 1-747; Length 740; Mismatches 0; Indels 0; Gaps 0;	
CC	Matches 14; Conservative	
DB	1229 gtagaggaagacta 1242	
DB	982 gtagaggaagacta 995	

ID	T89343	standard:	cDNA; 2063 BP.
AC	T89343;		
OT	11-MAR-1998	(first entry)	
OR	11-MAR-1998	Comment:	
PR	p62 cytoplasmic	T cell; B cell; development; activation; modulation;	
KW	cellular response; cell proliferation; autoimmune disease; etc.		
DE	Homo sapiens.		
TI			
FT			
FE	Key	Location/Qualifiers	
PF	CDS	6..1189	
PT		*****	
ST		/product= p62	
WT	W09722255-A1.		
PD	26-JUN-1997.	UI:004059.	
DP	13-DEC-1996.	US:174359.	
PA	[DAND] DANA FAREER CANCER INST INC.		
PI	Joung I, Shin J, Strominger JL, Vadlemludi RK;		
PR	WPJ: 97-3433/J.31.		
PS	SUBMITTER: p62 and p160 and corresponding proteins - used in the		
CC	treatment of autoimmune disease and for T and B cell proliferation,		
CT	e.g. for treatment of tumours		
CL	Claim 4: Fig 1: 175pp: English.		
FC	This cDNA sequence encodes a novel p62 cytoplasmic polypeptide which is		
CC	capable of interacting with a number of signalling molecules including heart, brain, placenta,		
CC	lung, liver, skeletal muscle, and thymus.		
CC	is capable of modulating T or B cell development and/or T or B cell		
CC	activation e.g. by modulation of Lck activity. It is also capable of		
CC	modulating degradation of cellular proteins e.g. cell cycle regulatory		
CC	proteins expressed in cell cycle dependent kinase inhibitors and arrestants.		
CC	to thereby modulate cell proliferation. As p62 acts to boost B cell		
CC	response and may be used to treat disorders where this is beneficial,		
CC			

Tue Nov 17 08:55:26 1998
 U9-08-887-977-9.rmg

CC	truncated rntin are cloned. The isolated nucleic acid and the	
CC	antibodies are used, in hybridization and immunosay, to screen	
CC	for HD.	
CC	Sequence 2085 BP: 746 A: 398 C: 557 G: 384 T;	
	Query Match	
	Best Local Similarity 100.0%; Score 15; DB 40; Length 2085;	
	Matches 15: Conservative 0; Pared 13; 63%+0;	
	Mismatches 0; Indels 0; Gaps 0;	
db	346 gctgagacactgctga 360	
py	599 gctgagacactgctga 613	

RESULT 311
ID Q25380 standard; cDNA: 2089 BP.
AC Q25380;
CC 22-NOV-1992 (first entry)
CC following acetolactate synthase (ALS) mutant 1 from clone 1B.
CC Acetolactate synthase herbicide resistant; maize; Indazolinone;
CC Inheritance: PDSUIT; as.
CC See maps.
CC NC N0206794.A.
CC NC N0206795.A.
CC NC N0206796.A.
CC PR 08-NOV-1991;
CC PR 14-NOV-1990; GS-Q24728.
CC PR (ICIL) IMPERIAL CHEM IND PLC.
CC bright SWZ, Chang MT, Evans JL, Macdonald MJ;
CC 11-12-2001/9/A.
CC CC transformed plants with resistance to herbicides, particu-
CC larly indazolinone cpds.
CC Claim 1: Fig 2A: 62pp; English.
CC Abstract [claiming] that authors were harvested from a field grown
CC maize plant containing a single copy of the wild type ALS gene
CC chemical mutagen ethyl methane sulphonate (EMS). The treated
CC pollen grains were brushed on to the silks of the detasseled inbred
CC female parent (coded Q253). The plants were grown to maturity and
CC then harvested. The M1 seed was sown and sprayed to be
CC "run-off" with herbicide. Plants were selected as exhibiting
CC tolerance of normally lethal dosages of the herbicide. Leaf
CC material from resistant plants was used to prepare cDNA libraries
CC which were screened to obtain DNA encoding the mutant ALS gene
CC sequence. When one of the clones was sequenced, it was found to
CC contain a single base mutation towards the 5' end of the coding
CC region of mutant 1 clone 1B (Q25380). This mutation, of G to A
CC transition, results in an Ala to Thr substitution. Such a mutation
CC has been reported to confer resistance to the herbicide acetola-
CC cynefluor herbicide but not conferring resistance to the herbicide
CC indazolinone. The sequence of mutant 1 clone 21 (Q25381) showed a single base

Tue Nov 17 08:55:26 1998

e.g. infections by pathogenic microorganisms, e.g. bacteria, viruses and protozoans) can be used to expand T cell populations for treating CC patients. However, it is important to select T cell populations that are able to render them resistant to HIV infection. Inhibitors of p24 can be used to reduce B or T cell responses and may be used to treat a variety of CC autoimmune diseases, e.g. diabetes mellitus, arthritis, multiple sclerosis, allergic reactions, Crohn's diseases etc.

SQ Sequence 2051 SP: 447 A: 588 C: 608 G: 440 T:
Query Match 1.3% Score 14; DB 36; Length 2083;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 1066 gsnatgatgagctc 1079
Cc 510 GAGGATGACTC 497

RESID	ID	V04262 standard; DNM, 2085 BP.	310
AC	V04262;		
CC	21-DEC-1997;		
CC	21-DEC-1998 (first entry)		
CC	21-DEC-1998 (second entry)		
CC	Restin; Hodgkin's disease; anti-		
KW	serological fishing; as:		
OS	Homo sapiens.	location/Qualifiers	
CC		2, 2085	
CC		/tag= a	
PPN	M09748721-Al.		
PPN	21-DEC-1997;		
PPN	21-DEC-1996; US-668128.		
PFP	(LUDWIG) LUDWIG INST CANCER RES.		
PFP	Friedrichshuh M.		
PPI	p-RPS28; w41586.		
PPI	Nucleic acid and derived proteins		
DPT	used in identifying immunoreact		
PPT	Claim 1, Page 33-35; 47pp; Engli		
PPT	(see W41586) that is truncated		
CCC	disruption of the alpha-helical		
CCC	the formation of filaments. No		
CCC	associated with this mutation		
CCC	isolated from biopsy tissues of		
CCC	designated 'aerologic' fishing		
CCC	markers of disease. Expression		
CCC	transfected cells; and antibody		

Tue Nov 17 08:55:26 1998

CC	mutation from the wild-type - A G to A transition 57 bases from the
CC	the coding region results in a Ser to Asp substitution in
CC	the HA sequence
80	Sequence 2089 BP; 433 A; 579 C; 579 G; 495 T;
	Query Match 1.3%; Score 34; DB 4; Length 2089;
	% Identical Similarity 8%; E-Value 6e-07;
Batches	14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB	7 cingulococcacagsty 22 ser_ccp1 ser_ccp2 ser_ccp3 ser_ccp4 ser_ccp5 ser_ccp6 ser_ccp7 ser_ccp8 ser_ccp9 ser_ccp10 ser_ccp11 ser_ccp12 ser_ccp13 ser_ccp14 ser_ccp15 ser_ccp16 ser_ccp17 ser_ccp18 ser_ccp19 ser_ccp20 ser_ccp21 ser_ccp22 ser_ccp23 ser_ccp24 ser_ccp25 ser_ccp26 ser_ccp27 ser_ccp28 ser_ccp29 ser_ccp30 ser_ccp31 ser_ccp32 ser_ccp33 ser_ccp34 ser_ccp35 ser_ccp36 ser_ccp37 ser_ccp38 ser_ccp39 ser_ccp40 ser_ccp41 ser_ccp42 ser_ccp43 ser_ccp44 ser_ccp45 ser_ccp46 ser_ccp47 ser_ccp48 ser_ccp49 ser_ccp50 ser_ccp51 ser_ccp52 ser_ccp53 ser_ccp54 ser_ccp55 ser_ccp56 ser_ccp57 ser_ccp58 ser_ccp59 ser_ccp60 ser_ccp61 ser_ccp62 ser_ccp63 ser_ccp64 ser_ccp65 ser_ccp66 ser_ccp67 ser_ccp68 ser_ccp69 ser_ccp70 ser_ccp71 ser_ccp72 ser_ccp73 ser_ccp74 ser_ccp75 ser_ccp76 ser_ccp77 ser_ccp78 ser_ccp79 ser_ccp80 ser_ccp81 ser_ccp82 ser_ccp83 ser_ccp84 ser_ccp85 ser_ccp86 ser_ccp87 ser_ccp88 ser_ccp89 ser_ccp90 ser_ccp91 ser_ccp92 ser_ccp93 ser_ccp94 ser_ccp95 ser_ccp96 ser_ccp97 ser_ccp98 ser_ccp99 ser_ccp100 ser_ccp101 ser_ccp102 ser_ccp103 ser_ccp104 ser_ccp105 ser_ccp106 ser_ccp107 ser_ccp108 ser_ccp109 ser_ccp110 ser_ccp111 ser_ccp112 ser_ccp113 ser_ccp114 ser_ccp115 ser_ccp116 ser_ccp117 ser_ccp118 ser_ccp119 ser_ccp120 ser_ccp121 ser_ccp122 ser_ccp123 ser_ccp124 ser_ccp125 ser_ccp126 ser_ccp127 ser_ccp128 ser_ccp129 ser_ccp130 ser_ccp131 ser_ccp132 ser_ccp133 ser_ccp134 ser_ccp135 ser_ccp136 ser_ccp137 ser_ccp138 ser_ccp139 ser_ccp140 ser_ccp141 ser_ccp142 ser_ccp143 ser_ccp144 ser_ccp145 ser_ccp146 ser_ccp147 ser_ccp148 ser_ccp149 ser_ccp150 ser_ccp151 ser_ccp152 ser_ccp153 ser_ccp154 ser_ccp155 ser_ccp156 ser_ccp157 ser_ccp158 ser_ccp159 ser_ccp160 ser_ccp161 ser_ccp162 ser_ccp163 ser_ccp164 ser_ccp165 ser_ccp166 ser_ccp167 ser_ccp168 ser_ccp169 ser_ccp170 ser_ccp171 ser_ccp172 ser_ccp173 ser_ccp174 ser_ccp175 ser_ccp176 ser_ccp177 ser_ccp178 ser_ccp179 ser_ccp180 ser_ccp181 ser_ccp182 ser_ccp183 ser_ccp184 ser_ccp185 ser_ccp186 ser_ccp187 ser_ccp188 ser_ccp189 ser_ccp190 ser_ccp191 ser_ccp192 ser_ccp193 ser_ccp194 ser_ccp195 ser_ccp196 ser_ccp197 ser_ccp198 ser_ccp199 ser_ccp200 ser_ccp201 ser_ccp202 ser_ccp203 ser_ccp204 ser_ccp205 ser_ccp206 ser_ccp207 ser_ccp208 ser_ccp209 ser_ccp210 ser_ccp211 ser_ccp212 ser_ccp213 ser_ccp214 ser_ccp215 ser_ccp216 ser_ccp217 ser_ccp218 ser_ccp219 ser_ccp220 ser_ccp221 ser_ccp222 ser_ccp223 ser_ccp224 ser_ccp225 ser_ccp226 ser_ccp227 ser_ccp228 ser_ccp229 ser_ccp230 ser_ccp231 ser_ccp232 ser_ccp233 ser_ccp234 ser_ccp235 ser_ccp236 ser_ccp237 ser_ccp238 ser_ccp239 ser_ccp240 ser_ccp241 ser_ccp242 ser_ccp243 ser_ccp244 ser_ccp245 ser_ccp246 ser_ccp247 ser_ccp248 ser_ccp249 ser_ccp250 ser_ccp251 ser_ccp252 ser_ccp253 ser_ccp254 ser_ccp255 ser_ccp256 ser_ccp257 ser_ccp258 ser_ccp259 ser_ccp260 ser_ccp261 ser_ccp262 ser_ccp263 ser_ccp264 ser_ccp265 ser_ccp266 ser_ccp267 ser_ccp268 ser_ccp269 ser_ccp270 ser_ccp271 ser_ccp272 ser_ccp273 ser_ccp274 ser_ccp275 ser_ccp276 ser_ccp277 ser_ccp278 ser_ccp279 ser_ccp280 ser_ccp281 ser_ccp282 ser_ccp283 ser_ccp284 ser_ccp285 ser_ccp286 ser_ccp287 ser_ccp288 ser_ccp289 ser_ccp290 ser_ccp291 ser_ccp292 ser_ccp293 ser_ccp294 ser_ccp295 ser_ccp296 ser_ccp297 ser_ccp298 ser_ccp299 ser_ccp300 ser_ccp301 ser_ccp302 ser_ccp303 ser_ccp304 ser_ccp305 ser_ccp306 ser_ccp307 ser_ccp308 ser_ccp309 ser_ccp310 ser_ccp311 ser_ccp312 ser_ccp313 ser_ccp314 ser_ccp315 ser_ccp316 ser_ccp317 ser_ccp318 ser_ccp319 ser_ccp320 ser_ccp321 ser_ccp322 ser_ccp323 ser_ccp324 ser_ccp325 ser_ccp326 ser_ccp327 ser_ccp328 ser_ccp329 ser_ccp330 ser_ccp331 ser_ccp332 ser_ccp333 ser_ccp334 ser_ccp335 ser_ccp336 ser_ccp337 ser_ccp338 ser_ccp339 ser_ccp340 ser_ccp341 ser_ccp342 ser_ccp343 ser_ccp344 ser_ccp345 ser_ccp346 ser_ccp347 ser_ccp348 ser_ccp349 ser_ccp350 ser_ccp351 ser_ccp352 ser_ccp353 ser_ccp354 ser_ccp355 ser_ccp356 ser_ccp357 ser_ccp358

```

RESULT 312
CO 23122 standard; DNA: 2094 BP.
AC Q23123;
DT 12-OCT-1992 (first entry)
DE ORF3 encoding Abi 105 phage resistance and partial ISS1 transposase
DE gene.
DE genbank accession: F00722.
DE insertion sequence: ISS1; transposase; resistance.
DE open reading frame: ORF 3; plasmid p1105; inverted repeat;
DE Streptococcus cremoris IL864.
OS Key Location/Qualifiers
CS
CDS
    /tag1 a
    /note="ORF3"
    /tag2 a
    /tag3 a
    /tag4 a
    /tag5 a
    /tag6 a
    /tag7 a
    /tag8 a
    /tag9 a
    /tag10 a
    /tag11 a
    /tag12 a
    /tag13 a
    /tag14 a
    /tag15 a
    /tag16 a
    /tag17 a
    /tag18 a
    /tag19 a
    /tag20 a
    /tag21 a
    /tag22 a
    /tag23 a
    /tag24 a
    /tag25 a
    /tag26 a
    /tag27 a
    /tag28 a
    /tag29 a
    /tag30 a
    /tag31 a
    /tag32 a
    /tag33 a
    /tag34 a
    /tag35 a
    /tag36 a
    /tag37 a
    /tag38 a
    /tag39 a
    /tag40 a
    /tag41 a
    /tag42 a
    /tag43 a
    /tag44 a
    /tag45 a
    /tag46 a
    /tag47 a
    /tag48 a
    /tag49 a
    /tag50 a
    /tag51 a
    /tag52 a
    /tag53 a
    /tag54 a
    /tag55 a
    /tag56 a
    /tag57 a
    /tag58 a
    /tag59 a
    /tag60 a
    /tag61 a
    /tag62 a
    /tag63 a
    /tag64 a
    /tag65 a
    /tag66 a
    /tag67 a
    /tag68 a
    /tag69 a
    /tag70 a
    /tag71 a
    /tag72 a
    /tag73 a
    /tag74 a
    /tag75 a
    /tag76 a
    /tag77 a
    /tag78 a
    /tag79 a
    /tag80 a
    /tag81 a
    /tag82 a
    /tag83 a
    /tag84 a
    /tag85 a
    /tag86 a
    /tag87 a
    /tag88 a
    /tag89 a
    /tag90 a
    /tag91 a
    /tag92 a
    /tag93 a
    /tag94 a
    /tag95 a
    /tag96 a
    /tag97 a
    /tag98 a
    /tag99 a
    /tag100 a
    /tag101 a
    /tag102 a
    /tag103 a
    /tag104 a
    /tag105 a
    /tag106 a
    /tag107 a
    /tag108 a
    /tag109 a
    /tag110 a
    /tag111 a
    /tag112 a
    /tag113 a
    /tag114 a
    /tag115 a
    /tag116 a
    /tag117 a
    /tag118 a
    /tag119 a
    /tag120 a
    /tag121 a
    /tag122 a
    /tag123 a
    /tag124 a
    /tag125 a
    /tag126 a
    /tag127 a
    /tag128 a
    /tag129 a
    /tag130 a
    /tag131 a
    /tag132 a
    /tag133 a
    /tag134 a
    /tag135 a
    /tag136 a
    /tag137 a
    /tag138 a
    /tag139 a
    /tag140 a
    /tag141 a
    /tag142 a
    /tag143 a
    /tag144 a
    /tag145 a
    /tag146 a
    /tag147 a
    /tag148 a
    /tag149 a
    /tag150 a
    /tag151 a
    /tag152 a
    /tag153 a
    /tag154 a
    /tag155 a
    /tag156 a
    /tag157 a
    /tag158 a
    /tag159 a
    /tag160 a
    /tag161 a
    /tag162 a
    /tag163 a
    /tag164 a
    /tag165 a
    /tag166 a
    /tag167 a
    /tag168 a
    /tag169 a
    /tag170 a
    /tag171 a
    /tag172 a
    /tag173 a
    /tag174 a
    /tag175 a
    /tag176 a
    /tag177 a
    /tag178 a
    /tag179 a
    /tag180 a
    /tag181 a
    /tag182 a
    /tag183 a
    /tag184 a
    /tag185 a
    /tag186 a
    /tag187 a
    /tag188 a
    /tag189 a
    /tag190 a
    /tag191 a
    /tag192 a
    /tag193 a
    /tag194 a
    /tag195 a
    /tag196 a
    /tag197 a
    /tag198 a
    /tag199 a
    /tag200 a
    /tag201 a
    /tag202 a
    /tag203 a
    /tag204 a
    /tag205 a
    /tag206 a
    /tag207 a
    /tag208 a
    /tag209 a
    /tag210 a
    /tag211 a
    /tag212 a
    /tag213 a
    /tag214 a
    /tag215 a
    /tag216 a
    /tag217 a
    /tag218 a
    /tag219 a
    /tag220 a
    /tag221 a
    /tag222 a
    /tag223 a
    /tag224 a
    /tag225 a
    /tag226 a
    /tag227 a
    /tag228 a
    /tag229 a
    /tag230 a
    /tag231 a
    /tag232 a
    /tag233 a
    /tag234 a
    /tag235 a
    /tag236 a
    /tag237 a
    /tag238 a
    /tag239 a
    /tag240 a
    /tag241 a
    /tag242 a
    /tag243 a
    /tag244 a
    /tag245 a
    /tag246 a
    /tag247 a
    /tag248 a
    /tag249 a
    /tag250 a
    /tag251 a
    /tag252 a
    /tag253 a
    /tag254 a
    /tag255 a
    /tag256 a
    /tag257 a
    /tag258 a
    /tag259 a
    /tag260 a
    /tag261 a
    /tag262 a
    /tag263 a
    /tag264 a
    /tag265 a
    /tag266 a
    /tag267 a
    /tag268 a
    /tag269 a
    /tag270 a
    /tag271 a
    /tag272 a
    /tag273 a
    /tag274 a
    /tag275 a
    /tag276 a
    /tag277 a
    /tag278 a
    /tag279 a
    /tag280 a
    /tag281 a
    /tag282 a
    /tag283 a
    /tag284 a
    /tag285 a
    /tag286 a
    /tag287 a
    /tag288 a
    /tag289 a
    /tag290 a
    /tag291 a
    /tag292 a
    /tag293 a
    /tag294 a
    /tag295 a
    /tag296 a
    /tag297 a
    /tag298 a
    /tag299 a
    /tag300 a
    /tag301 a
    /tag302 a
    /tag303 a
    /tag304 a
    /tag305 a
    /tag306 a
    /tag307 a
    /tag308 a
    /tag309 a
    /tag310 a
    /tag311 a
    /tag312 a
    /tag313 a
    /tag314 a
    /tag315 a
    /tag316 a
    /tag317 a
    /tag318 a
    /tag319 a
    /tag320 a
    /tag321 a
    /tag322 a
    /tag323 a
    /tag324 a
    /tag325 a
    /tag326 a
    /tag327 a
    /tag328 a
    /tag329 a
    /tag330 a
    /tag331 a
    /tag332 a
    /tag333 a
    /tag334 a
    /tag335 a
    /tag336 a
    /tag337 a
    /tag338 a
    /tag339 a
    /tag340 a
    /tag341 a
    /tag342 a
    /tag343 a
    /tag344 a
    /tag345 a
    /tag346 a
    /tag347 a
    /tag348 a
    /tag349 a
    /tag350 a
    /tag351 a
    /tag352 a
    /tag353 a
    /tag354 a
    /tag355 a
    /tag356 a
    /tag357 a
    /tag358 a
    /tag359 a
    /tag360 a
    /tag361 a
    /tag362 a
    /tag363 a
    /tag364 a
    /tag365 a
    /tag366 a
    /tag367 a
    /tag368 a
    /tag369 a
    /tag370 a
    /tag371 a
    /tag372 a
    /tag373 a
    /tag374 a
    /tag375 a
    /tag376 a
    /tag377 a
    /tag378 a
    /tag379 a
    /tag380 a
    /tag381 a
    /tag382 a
    /tag383 a
    /tag384 a
    /tag385 a
    /tag386 a
    /tag387 a
    /tag388 a
    /tag389 a
    /tag390 a
    /tag391 a
    /tag392 a
    /tag393 a
    /tag394 a
    /tag395 a
    /tag396 a
    /tag397 a
    /tag398 a
    /tag399 a
    /tag400 a
    /tag401 a
    /tag402 a
    /tag403 a
    /tag404 a
    /tag405 a
    /tag406 a
    /tag407 a
    /tag408 a
    /tag409 a
    /tag410 a
    /tag411 a
    /tag412 a
    /tag413 a
    /tag414 a
    /tag415 a
    /tag416 a
    /tag417 a
    /tag418 a
    /tag419 a
    /tag420 a
    /tag421 a
    /tag422 a
    /tag423 a
    /tag424 a
    /tag425 a
    /tag426 a
    /tag427 a
    /tag428 a
    /tag429 a
    /tag430 a
    /tag431 a
    /tag432 a
    /tag433 a
    /tag434 a
    /tag435 a
    /tag436 a
    /tag437 a
    /tag438 a
    /tag439 a
    /tag440 a
    /tag441 a
    /tag442 a
    /tag443 a
    /tag444 a
    /tag445 a
    /tag446 a
    /tag447 a
    /tag448 a
    /tag449 a
    /tag450 a
    /tag451 a
    /tag452 a
    /tag453 a
    /tag454 a
    /tag455 a
    /tag456 a
    /tag457 a
    /tag458 a
    /tag459 a
    /tag460 a
    /tag461 a
    /tag462 a
    /tag463 a
    /tag464 a
    /tag465 a
    /tag466 a
    /tag467 a
    /tag468 a
    /tag469 a
    /tag470 a
    /tag471 a
    /tag472 a
    /tag473 a
    /tag474 a
    /tag475 a
    /tag476 a
    /tag477 a
    /tag478 a
    /tag479 a
    /tag480 a
    /tag481 a
    /tag482 a
    /tag483 a
    /tag484 a
    /tag485 a
    /tag486 a
    /tag487 a
    /tag488 a
    /tag489 a
    /tag490 a
    /tag491 a
    /tag492 a
    /tag493 a
    /tag494 a
    /tag495 a
    /tag496 a
    /tag497 a
    /tag498 a
    /tag499 a
    /tag500 a
    /tag501 a
    /tag502 a
    /tag503 a
    /tag504 a
    /tag505 a
   
```

CC 10th were ligated into the BamHI site of pTZ192. One recombinant
 CC plasmid contained the insert coding for the ADL 105 nucleomem.
 CC This plasmid was designated pTZ193. A 700bp EcoRV-XbaI fragment
 CC was excised from pTZ193 and the religated plasmid, containing a
 CC 5.7kb insert, was designated pTZ194. The 5.7kb insert was partially
 CC digested with EcoRV and the resulting fragments were ligated into
 CC the transposon from insertion sequence IS91. The sequence
 CC given here is that of ORF3, which codes for a protein involved in
 CC the bacteriophage resistance mechanism, part of the IS91
 CC transposase gene and one of the inverted repeat (IS) sequences
 CC of the transposon. See Q3321, Q3322, Q3323, Q3324, Q3325, Q3326,
 SQ Sequence 2094 BP; 794 A; 332 C; 376 G; 602 T;

Query Match 1.34; Score 14; DB 4; Length 2094;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 586 aaagaacacaaaga 599
 Cp 645 AAGAACCAACAAAGA 632

RESULT 313
 ID T36141 standard; cDNA; 2096 BP.
 DT 23-DEC-1995 (first entry)
 DE Human L-asparaginase cDNA.
 KW L-asparaginase; antitumour; cancer; lymphoma; leukaemia; therapy;
 SS.
 FH Homo sapiens. Location/Qualifiers
 FT cds 93..1814
 FT /tag= a
 FT EP-726131-A2.
 FT P-PSDB: W47414.
 PP 07-FEB-1995; 300803
 PP 08-FEB-1995; JP-042564.
 PA (HAYB) HAYASHIBARA SEIITSU KAGAKU.
 PI Ario T., Kurimoto M., Tanai M., Torigoe K.
 DR P-PSDB: W47414.
 DR P-PSDB: W47414.
 PT DNA encoding mammalian L-asparaginase - for prodn. of recombinant
 PT enzyme for tumour therapy
 PS Claim 7; Page 34-36; 37pp; English.
 CC The present sequence encodes the wild type human L-asparaginase.
 CC A mammalian polypeptide having L-asparaginase activity, can be used
 CC to treat malignant tumours, leukaemia and lymphoma.
 CC host (e.g. E. coli) cells.
 SQ Sequence 2096 BP; 397 A; 633 C; 685 G; 361 T;

Query Match 1.34; Score 14; DB 22; Length 2096;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 339 ttgactctcaatga 352
 Cp 6 ttgactctcaatga 19

RESULT 314
 ID T36141 standard; cDNA to mRNA; 2096 BP.
 DT 05-JUN-1998 (first entry)
 DE cDNA for wild type human L-asparaginase.
 KW L-asparaginase; treatment; malignant tumour;
 OS Leukemia; lymphoma; human; ds.
 FH Key captions. Location/Qualifiers
 FT cds 93..1814
 FT /tag= a
 FT EP-811687-A2.
 FT P-PSDB: W47414.
 PP 06-JUN-1997; 303896.
 PP 07-JUN-1996; JP-168172.
 PA (Kato) KATOH SHINJIRO KAGAKU.
 PI Ario T., Kurimoto M., Tanai M., Yumoto K.
 DR P-PSDB: W47414.
 DR P-PSDB: W47414.
 PT Mammalian L-asparaginase polypeptide - useful for treating cancer
 PT The present sequence encodes the wild type human L-asparaginase.
 CC A mammalian polypeptide having L-asparaginase activity, can be used
 CC to treat malignant tumours, leukaemia and lymphoma.
 SQ Sequence 2096 BP; 397 A; 633 C; 685 G; 361 T;

Query Match 1.34; Score 14; DB 39; Length 2096;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 339 ttgactctcaatga 352
 Cp 6 ttgactctcaatga 19

RESULT 315
 ID Q05676 standard; cDNA; 2100 BP.
 DT 03-JAN-1991 (first entry)
 DE PA-gamma.
 KW Thrombolytic agent; fibrinolysis; saliva; vampire bat; ss.

OS Desmodus rotundus.
 FH Key Location/Qualifiers
 FT cds 180..1364
 FT /tag= a
 FT /product=v-PA_gamma
 FT EP-383417-A.
 FT P-PSDB: R06486.
 PP 22-AUG-1990. 300643
 PP 11-FEB-1989; DE-904580.
 PP 30-MAY-1989; DE-917949.
 PA (SCHD) SCHERING AG.
 PI Balduz B., Donner P., Schleuning WD., Alagon A., Boidol W.
 DR P-PSDB: R06486.
 DR P-PSDB: R06486.
 PT New fibrin specific thrombolytic agent v-PA - isolated from
 PT saliva of Desmodus sp. bats.
 CC The v-PA was obtained by expression of the gene in a
 CC cDNA library prep. from RNA from the salivary glands of vampire
 CC bats. The v-PA protein produced on expression of the gene is a
 CC low molecular form consisting of a kringle domain and a protease
 CC domain. The v-PA is highly specific and dissolves blood clots.
 CC See also Q05673-Q05677.
 SQ Sequence 2100 BP; 598 A; 512 C; 521 G; 469 T;

Query Match 1.34; Score 15; DB 1; Length 2100;
 Best Local Similarity 100.0%; Pred. No. 8.40e+00;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 489 tcaaaacttggtgc 503
 Cp 686 tcaaaacttggtgc 700

RESULT 316
 ID T36141 standard; DNA; 2107 BP.
 DT 20-MAY-1998 (first entry)
 DE Human steroid 5-alpha reductase cDNA.
 KW Steroid 5-alpha reductase coding sequence.
 KW Steroid 5-alpha reductase; enzyme; testosterone conversion; inhibitor;
 OS human; ss.
 FH Homo sapiens. Location/Qualifiers
 FT cds 31..810
 FT /tag= a
 FT EP-8567921-A.
 FT P-PSDB: W47414.
 PP 01-JUN-1995; 457616.
 PP 18-NOV-1991; US-794859.
 PA (TEKA) UNIV TEXAS SYSTEM.
 PI Anderson S., Russell DW;

PR 01-JUN-1995; US-457616.
 PA (TEKA) UNIV TEXAS SYSTEM.
 PI Anderson S., Russell DW.
 DR P-PSDB: W47414.
 DR P-PSDB: W47414.
 PT Production of recombinant steroid 5-alpha reductase enzyme - by
 PT culturing cell containing DNA encoding the enzyme
 PS Disclosure; Columns 45-50; 70pp; English.
 CC The present sequence encodes the wild type human steroid 5-alpha reductase.
 CC The encoded enzyme can be produced by the method of the invention. The
 CC method is for producing a steroid 5-alpha reductase, and comprises
 CC preparing a recombinant host cell containing a DNA segment encoding a
 CC steroid 5-alpha reductase and culturing the cell under conditions such
 CC that the cell produces and secretes the steroid 5-alpha reductase.
 CC Salpa reductase produced by the method is used for identifying
 CC substances that affect the enzymatic activity of steroid
 CC 5-alpha reductase. Substances identified as inhibiting steroid
 CC 5-alpha reductase activity can be used for inhibiting the conversion of
 CC cholesterol to androgens.
 SQ Sequence 2107 BP; 475 A; 482 C; 486 G; 664 T;

Query Match 1.34; Score 14; DB 37; Length 2107;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 379 tgaacatggagat 392
 Cp 419 tgaacatggagat 406

RESULT 317
 ID Q97380 standard; DNA; 2107 BP.
 DT 01-APR-1996 (first entry)
 DE Human type I steroid 5-alpha reductase cDNA.
 KW Steroid 5-alpha reductase; sexual development; differentiation;
 KW probe; recombinant; inhibit; prostatic hyperplasia; acne; hirsutism;
 OS Homo sapiens. Location/Qualifiers
 FH Key 31..810
 FT cds 31..810
 FT /tag= a
 FT /product= steroid_5-alpha_reductase_type-I
 US5422262-A.
 PP 06-JUN-1995; 517651.
 PP 30-APR-1990; 517651.
 PP 18-NOV-1991; US-795959.
 PA (TEKA) UNIV TEXAS SYSTEM.
 PI Anderson S., Russell DW;

NR 1: 95-214659/28.
 CC Steroid 5 alpha-reductase nucleic acid segments and recombinant
 PT vectors, where the sequences are useful in e.g. analysis of normal
 PT and abnormal sexual differentiation.
 PS Claim 3: Column 47-52: 79pp; English.
 CC dihydroxyacetone. The enzyme catalyze the conversion of testosterone to
 CC dihydroxyacetone. The enzyme has been isolated and purified. The human enzyme
 CC cDNA sequence (G97180) has been isolated and purified. The human enzyme
 CC cDNA can be used in the prep. of genetic constructs for the large
 CC scale production of SHD5A or as probes for enzyme-encoding sequences
 CC of normal and abnormal sexual differentiation. The enzyme is useful in the analysis
 CC of normal and abnormal sexual differentiation. The enzyme is useful in the analysis
 CC of hyperplasia, male pattern baldness; acne; hirsutism and endometriosis.
 SQ Sequence 2107 BP: 475 A; 482 C; 486 G; 664 T;
 Query Match 1.3%; Score 14; DB 17; Length 2107;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 379 tctgactctgcgc 392
 Cp 419 tctgactctgcgc 406

RESULT 318
 ID Q14387 standard; DNA: 2107 BP.
 AC Q14387;
 DT 10-FEB-1992 (first entry)
 DE Human steroid 5-alpha reductase gene.
 KW Human steroid 5-alpha reductase; gene; hirsutism; dihydroxyacetone; ss.
 OS Homo sapiens.
 FN Key Location/Qualifiers
 FT cds 31..810
 FT /tag= a
 FT /tag= b
 FT polyA_signal 2582..2597
 FT /tag= a
 FT /tag= b
 FT W02117251-A.
 PD 24-NOV-1991.
 PD 30-APR-1990; US-517661.
 PA (TEMA) UNIV TEXAS SYST.
 PI Anderson S, Russell DW;
 DR P-5039; R3102/48.
 DT New DNA encoding human steroid 5 alpha-reductase - and corresp.
 PT vectors and host cells, useful to screen potential inhibitors for
 PT treatment of endocrine abnormalities, etc.
 PS Claim; Fig 1; 84pp; English.
 CC Rat cDNA probes (see Q14386) were used to screen a human prostate

CC cDNA library, to identify a 2.1 kb fragment encoding a steroid
 CC 5-alpha reductase of 259 amino acids. The invention also covers
 CC fragments of the human-derived sequence as well as considering
 CC fragments of the human-derived sequence as well as considering
 SQ Sequence 2107 BP: 475 A; 482 C; 486 G; 664 T;
 Query Match 1.3%; Score 14; DB 3; Length 2107;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 379 tctgactctgcgc 392
 Cp 419 tctgactctgcgc 406

RESULT 319
 ID Q14202 standard; cDNA: 2109 BP.
 AC Q14202;
 DT 27-APR-1998 (first entry)
 DE Human calpain cDNA.
 KW Calpain; human; leukocyte; calcium dependent cysteine protease;
 KW screening; activator; inhibitor; treatment; prevention; cancer;
 KW Alzheimer's disease; amyotrophy; cataracts; collagen disease;
 KW ischaemic heart disease; atherosclerosis; arthritis; da.
 OS Homo sapiens.
 FN Key Location/Qualifiers
 FT cds 1..2109
 FT /tag= a
 FT /tag= b
 FT /product= calpain
 FT /note= "stop codon not given"
 DT BP-709892-43.
 DT 08-OCT-1997.
 PP 03-APR-1997; 105308.
 PP 05-APR-1996; JP-083649.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR Takeda Pharmaceutical Co., Ltd., A. Shimada, Y;
 DR WPI: 97-483674/45.
 DR P-PS29; R41564.
 DT Human calpain protein and related DNA - useful for drug screening
 PT and treating cancer, stroke, etc.
 PS The present sequence encodes calpain.
 CC The present sequence encodes calpain, a human leukocyte derived
 CC calcium dependent cysteine protease. Calpain can be used to screen
 CC for compounds that activate or inhibit its proteolytic activity.
 CC Calpain DNA can be used to treat or prevent cancer, cerebral
 CC Alzheimer's disease, amyotrophy, cataracts, ischaemic heart
 CC disease, atherosclerosis, arthritis or collagen disease.
 SQ Sequence 2109 BP: 499 A; 529 C; 623 G; 458 T;

Query Match 1.3%; Score 14; DB 37; Length 2109;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 797 tctgactctgcgc 810
 Cp 19 tctgactctgcgc 6

RESULT 320
 ID Q56794 standard; DNA: 2124 BP.
 AC Q56794;
 DT 27-APR-1994 (first entry)
 DE Granulin; keratinocytes; wound healing; inhibition; peptide;
 KW granulocytes; leucocytes; ss.
 OS Homo sapiens.
 FN Key Location/Qualifiers
 FT misc_difference 1374..1801
 FT /tag= a
 FT /transl_except= ATG encodes Valine.
 FT misc_difference 1640..1642
 FT /tag= a
 FT /transl_except= CAG encodes Glycine.
 W09315195-A.
 PD 05-AUG-1993.
 PD 03-FEB-1992; GB-099233.
 PA (SOL2) SOLOMON S.
 DR WPI: 93-32028/40.
 DT New cysteine rich granulin peptide(s) from leucocyte(s) - are
 PT keratinocyte inhibitors useful topically for wound healing
 PS Disclosure: Figure 4c; 53pp; English.
 CC The granulin inhibits keratinocytes and is useful in formulations
 CC for the treatment of wound healing.
 SQ Sequence 2124 BP: 385 A; 685 C; 630 G; 426 T;
 Query Match 1.3%; Score 14; DB 9; Length 2124;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 754 tctgactctgcgc 767
 QY 691 tctgactctgcgc 904

RESULT 321
 ID Q63972 standard; cDNA: 2139 BP.
 AC Q63972;
 DT 03-APR-1993 (first entry)
 DE Sequence of clone pVAM09R3 which contains probes for the in vitro
 DT diagnosis of patients presenting an alteration on their chromosome
 KW PCR primer: Charcot-Marie-Tooth type 1 disease; in vitro diagnosis;
 OS Homo sapiens.

DT 08-DEC-1994 (first entry)
 DE Megakaryocyte potentiator (pMP027).
 KW Megakaryocyte potentiator; Meg-Pot; thrombocytopenia;
 KW platelet; ss.
 OS Homo sapiens.
 FN Key Location/Qualifiers
 FT cds 57..1965
 FT /tag= a
 FT /product= Meg-Pot
 FT /tag= b
 FT /tag= c
 FT /note= "Claim 1; 594 amino acid product"
 FT misc_rna 211..954
 FT /tag= c
 FT /note= "Claim 6; 248 amino acid product"
 W09410312-A.
 PD 11-MAY-1994.
 PP 25-OCT-1993; J01540.
 PP 11-OCT-1992; JP-101387.
 PP 11-NOV-1992; JP-101387.
 PP 09-DEC-1992; JP-329546.
 PA (CHUS) CHUGAI SEITAKU KK.
 PI Hattori K, Koijima T, Oh-eda M, Yamaguchi N;
 DR P-PS29; R51992.
 DT Rev mega.karyocyte potentiator - for potential treatment of
 PT thrombocytopenia
 PS Disclosure: Page 52-57; 74pp; Japanese.
 CC The present sequence encodes a protein which has potential
 CC use in treatment of thrombocytopenia and low platelet function.
 CC Sequence 2129 BP: 389 A; 715 C; 658 G; 367 T;
 Query Match 1.3%; Score 14; DB 11; Length 2129;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 878 ggcgactctgcgc 891
 Cp 376 ggcgactctgcgc 363

RESULT 322
 ID Q05222 standard; DNA: 2133 BP.
 AC Q05222;
 DT 03-APR-1993 (first entry)
 DE Sequence of clone pVAM09R3 which contains probes for the in vitro
 DT diagnosis of patients presenting an alteration on their chromosome
 KW PCR primer: Charcot-Marie-Tooth type 1 disease; in vitro diagnosis;
 OS Homo sapiens.

KW Barley; heat resistant; promoter: genetic engineering; transgenic;
 PL Plant; improved maturation; da.
 PS WO9702133-A.
 PD 23-JAN-1997.
 PF 05-JUL-1996; J01866.
 PR 05-JUL-1996; JP-191028.
 PT Ito K, Kihara M, Ohada T, Yoshigi N;
 DR WO9702133-A.
 DT IPI; 97-108966/20.
 CC Expression promoter for genes inserted into barley seeds.
 CC Heat-resistant beta-amylase gene inserted into barley seeds.
 CC Heat-resistant beta-amylase gene inserted into barley seeds.
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences
 CC used for integration of this promoter into a reporter plasmid. The
 CC plasmid is used for the production of transgenic barley plants.
 CC Examples: Column 115-120: 86pp; English.
 SQ Sequence 2142 BP; 614 T; 395 C; 482 G; 482 A; 482 T; 614 T.

Query Match 1.34; Score 14; DB 33; Length 2142;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 327 attcttcctccca 340
 Cp 994 ATTCTTCTCTGAC 991

RESULT 329
 ID T42218 standard; cDNA; 2152 BP.
 AC T78594.
 DT 08-OCT-1997 (first entry)
 DE Human TATA-binding protein associated factor, hTAP1100 cDNA.
 KW TATA-binding protein associated factor; TAP; nuclear protein;
 KW RNA polymerase II; transcription; messenger RNA; nuclear fraction;
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FE Key 1..2115
 FT cds /note= "no start codon"
 FT 1..2115
 PD US5637686-A.
 PD 10-JUN-1997.
 PR 28-JAN-1993; 013412.
 PR 30-JUN-1993; US-087119.
 PR 28-JAN-1994; US-188582.
 PR 09-MAY-1996; US-646715.
 PA (REGC) UNIV CALIFORNIA.
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanase N;
 DR P-PSDB; W06085.

PT in screening, diagnostics and therapeutics.
 PS Disclosure: Page 143-147; 180pp; English.
 CC The TATA-binding protein associated factor hTAP1100 (including
 CC diagnostic and therapeutic products) are used in drug screening,
 CC diagnosis and therapeutic products in vitro.
 CC specific biochemical assays for screening compounds that agonize or
 CC antagonize selected transcription factors involved in regulating
 CC gene expression associated with human pathology.
 SQ Sequence 2152 BP; 627 A; 451 C; 534 G; 539 T;

Query Match 1.34; Score 15; DB 12; Length 2152;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1505 attattttgtctag 1519
 Cy 22 ATTATTGTCTGAC 36

RESULT 329
 ID T42218 standard; cDNA; 2152 BP.
 AC T78594.
 DT 22-JUN-1994 (first entry)
 DE Human TATA-binding protein; TBP associated factor hTAP1100 gene.
 KW RNA polymerase II; transcription; messenger RNA; nuclear fraction;
 KW TBP associated factor (TAP) designated TAP1100.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FE Key 1..2115
 FT cds /product= "human TAP1100"
 FT 1..2115
 PD US534410-A.
 PD 09-JUL-1996.
 PR 28-JAN-1993; 013412.
 PR 28-JAN-1993; US-013412.
 PR 30-JUN-1993; US-087119.
 PR 28-JAN-1994; US-188582.
 PA (REGC) UNIV CALIFORNIA.
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanase N;
 DR P-PSDB; W06085.
 PT Screen for cDNA that bind human TATA-binding protein associated
 PT factor - by testing ability to bind to polypeptide fragments of the
 PT factor, useful as (antagonists of transcription factors involved in
 PT transcription).
 PS Examples: Column 115-120: 86pp; English.
 CC This is the nucleotide sequence encoding the human TATA-binding protein
 CC (TBP) associated factor (TAP) designated TAP1100. The protein is a
 CC component of the TFIID fraction required for reconstituting RNA

PI Tjian R, Wang E, Weinsierl ROJ;
 DR WPI; 97-31913/29.
 DR P-PSDB; W25018.
 DT 22-JUN-1994 (first entry)
 DE Nucleic acids encoding human TATA-binding protein associated factor
 DE for modulating transcription of TAPs.
 PS Claim 11: Column 115-120: 86pp; English.
 CC T79594 encodes human TATA-binding protein associated factor (TAP)
 CC polypeptide. hTAP1100 cDNA weight 100kD. TAP peptides
 CC hTAP1100, hTAP1100, hTAP1100, hTAP1100, hTAP1100, hTAP1100,
 CC hTAP1130, hTAP1250, hTAP248 and hTAP1110 and nucleic acids encoding
 CC them, are used to modulate transcription, including transcription
 CC initiation. TAPs are nuclear proteins involved in RNA polymerase I,
 CC RNA polymerase II, and RNA polymerase III transcription. TAPs are
 CC TAP, an activator, or TBP (TATA-binding protein) competitively
 CC inhibiting association of a TAP domain with another compound, typically
 CC a protein like TBP or another TAP, an activator, or DNA.
 SQ Sequence 2152 BP; 627 A; 449 C; 536 G; 540 T;

Query Match 1.34; Score 15; DB 32; Length 2152;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1505 attattttgtctag 1519
 Cy 22 ATTATTGTCTGAC 36

RESULT 328
 ID Q70732 standard; cDNA; 2152 BP.
 AC Q70732.
 DT 22-JUN-1994 (first entry)
 DE TATA-binding protein associated factor hTAP1100 cDNA.
 KW TATA-binding protein associated factor; hTAP1100; as; screening;
 KW diagnostic; therapeutic; gene transcription regulation.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FE Key 1..2115
 FT cds /tag= a
 FT 1..2115
 PD WO9417087-A.
 PD 04-AUG-1994.
 PR 28-JAN-1993; US-013412.
 PR 30-JUN-1993; US-087119.
 PA (REGC) UNIV CALIFORNIA.
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanase N;
 DR WPI; 94-264019/32.
 DR P-PSDB; R56495.
 PT TATA-binding protein associated protein factors - and
 PT corresponding nucleotide sequence and deriv. antibodies, useful

CC polymerase II in vitro transcription activity. The encoded protein has
 CC an estimated mol. wt. of 100 kD by SDS-PAGE.
 CC The invention relates to purified proteins involved in transcription
 CC by RNA polymerase II, the RNA polymerase which transcribes messenger
 CC RNA. RNA polymerase II transcription products in vitro upon addition
 CC of purified TAPs, hTAP1100, hTAP1100, hTAP1100, hTAP1100, hTAP1100, hTAP1100,
 CC to RNA polymerase II holoenzyme. Fraction TFIID has been shown to
 CC contain a TAP and other TAPs. Purification of TFIID and separation of
 CC its components reveals 7 proteins ranging in size from 30-350 kD.
 CC The invention relates to the cloning of the genes from lambda-phage
 CC expression libraries.
 SQ Sequence 2152 BP; 627 A; 449 C; 536 G; 540 T;

Query Match 1.34; Score 15; DB 25; Length 2152;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1505 attattttgtctag 1519
 Cy 22 ATTATTGTCTGAC 36

RESULT 310
 ID T42218 standard; cDNA; 2152 BP.
 AC Q54742.
 DT 22-JUN-1994 (first entry)
 DE Toxoplasma GP28.5 antigen coding sequence.
 KW Toxoplasma; Toxoplasma; tachyzoite; excretion-secretion antigen;
 KW dense granule; as; glycoprotein; parasitophorous vacuole;
 OS Toxoplasma gondii.
 FH Key Location/Qualifiers
 FE Key 986..1035
 FT exon 1036..1274
 FT intron 1275..1682
 FT exon /tag= b
 FT 1275..1682 /tag= c
 PD WO9125689-A.
 PD 21-DEC-1993.
 PR 15-JUN-1993; F00575.
 PR 15-JUN-1993; FR-007206.
 PA (INSP) INST PASTEUR LILLE.
 PA (INSP) INST PASTEUR LILLE.
 PI Capron A, Cesbron M, Maes P, Mercier C, Tartar A;
 DR WPI; 94-007145/01.
 PT New nucleic acid for Toxoplasma gondii GP 28.5 antigen - for
 PT deriving proteins and peptide(s) and useful in vaccines and
 PT diagnosis, vectors and transformed cells

S Claim 2: Page 23-25; 37pp; French.
A T-pendil genomic library in EMBL3 was screened with a fragment of
a clone isolated by immunoscreening a T-pendil strain RR cDNA
library with a T-pendil strain RR cDNA library. The clone that
encodes the complete 38 kDa antigen was identified. The antigen is
the principal secretion-activation antigen of T-pendil tachyzoites
and represents a good candidate for effective vaccine production.
Sequence 212 BP, 555 A, 335 C, 397 T.

Query Match 1.38; Score 14; DB 9; Length 212;
Local Similarity 100.0%; Read No. 3,746,02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

b 1429 aaagagcgaaggtg 1442
193 aaagagcgaaggtg 180

D D64125 standard; DNA: 2154 BP.
C O3-FEB-1995 (first entry) EBV-L1 gene
T C064125:
C O3-FEB-1995 (last entry) EBV-L1 gene
W W034125-19-A: EBV infection detection; diagnosis;
W Epstein Barr virus; EBV induction; detection; diagnosis;
W lymphocytes; antigen; growth; differentiation; mediator;
W infectious mononucleosis; ss.
W homo sapiens. Location/Qualifiers
T cds
T 64..1200
T /tag= a
T /product= Epstein Barr virus induced polypeptide.
D W09412519-A.
D 09-JUN-1994.
D 08-OCT-1993: U09636.
P 25-NOV-1992: US-980518.
R BUCKINGHAM MEMORIAL HOSPITAL.
R BUCKINGHAM WAY KINGS LANCERS
R WPI: 94-200183/24.
R P-PSDB: AS0709.
R DNA coding for Epstein Barr Virus induced (EBI) polypeptide(s)
R protein product of EBV-L1 gene.
S hydrolysis or by immunosay
T Claim 2: Page 52-54; 8pp; English.
C
C EBV infected B lymphocytes recapitulate features of antigen
C activation antigens and adhesion molecules secreting Ig and
C proliferation. Unlike antigen stimulated B lymphocytes, EBV
C infected B lymphocytes continue to proliferate (in vitro) as
C long as EBV and antigen are present. These results suggest
C effects of EBV and antigen on EBV induced species likely to include
C mediators of antigen induced B lymphocyte growth or differentiation.

```

SQ Sequence      2154 BP; 486 A;   647 C;   543 G;   478 T;
Query Match     1.1%; Score 14; DB 11; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3,47+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Db 370 gongocattctt 393
    |||||
    |||||
QY 244 CGAGGACGGCTGT 257

RESULT 332
ID O66160 standard; cDNA; 2160 BP.
IC 100%
PC O-PS-1995 (first entry)
DE Putative seven transmembrane receptor (V1-B) coding sequence (cDNA).
DS Primer: seven transmembrane receptor; receptor: amplification; PCR;
KW polymerase chain reaction; ss.
OS Homo sapiens.
GN 147
NC FT
CD 5'
        Location/Qualifiers
            /locus_tag =
            /product= Seven transmembrane receptor V31.

```

PP 0 9-Nov-1993; U11153
 PP 17-MOV-1992; US-97745Z.
 PA (ICOS-) ICOS CORP.
 PI Godiska R., 1008 Gray FW, Schwickart VL;
 PZ P-PMDS: 8537474.
 DR DNA encoding seven trans-membrane receptors - used to develop
 PT proteins for use as therapeutic or diagnostic agents for conditions
 PT involving the receptor; 10pp; English.
 CC A human cDNA encoding the seven transmembrane receptor Vsl was
 CC isolated by first amplifying a partial cDNA clone from a human
 CC tonsil cDNA library using two primers (O66150, O66155). The resulting
 CC sequences (O66150, O66157) showed a hybridizing band was isolated from
 CC gels and cloned. The resulting clone was named pvl-5'end (O66158). A
 CC full length clone was isolated from a peripheral blood
 CC monocyte cell library using Vsl specific primers (O66159, O66160).
 CC Cloned PCR products were sequenced. The vsl cDNA insert in the clone was
 CC designated cDNA vsl-b (O66160).
 QU Query sequence 2160 BP: 645 C: 546 G: 479 T:
 133A: Score 14; DP 11; Length 2160;
 Match Local Similarity 100.0%; Pres. No. 3,477+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue Nov 17 08:55:26 1998

Tue Nov 17 08:55:26 1998

Page 332

b	370	GAGACATCTCTT	383
c			
d			
e			
f			
g			
h			
i			
j			
k			
l			
m			
n			
o			
p			
q			
r			
s			
t			
u			
v			
w			
x			
y	244	GAGACATCTCTT	257
z			
A			
B			
C			
D			
E			
F			
G			
H			
I			
J			
K			
L			
M			
N			
O			
P			
Q			
R			
S			
T			
U			
V			
W			
X			
Y			
Z			
[
\			
]			
_			
`			
a			
b			
c			
d			
e			
f			
g			
h			
i			
j			
k			
l			
m			
n			
o			
p			
q			
r			
s			
t			
u			
v			
w			
x			
y			
z			
A			
B			
C			
D			
E			
F			
G			
H			
I			
J			
K			
L			
M			
N			
O			
P			
Q			
R			
S			
T			
U			
V			
W			
X			
Y			
Z			
[
\			
]			
_			
`			
a			
b			
c			
d			
e			
f			
g			
h			
i			
j			
k			
l			
m			
n			
o			
p			
q			
r			
s			
t			
u			
v			
w			
x			
y			
z			
A			
B			
C			
D			
E			
F			
G			
H			
I			
J			
K			

```

CC Fluorimethyl arthritis, pericarditis and coronary atherosclerosis
CC increase tubular regeneration of kidney cells, e.g. following acute
CC renal failure.
CC Sequence 2162 BP: 607 A: 573 C: 571 G: 411 T:
CC
CC Query Match
CC 1.33: Score 14; DB 39; Length 2162;
CC Best Local Similarity 100.0%; Pred. No. 3,47e+02;
CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC BD 1642 tectoglycoctect 1655
CC
CC 11111111111111111111
CC 73 TACTGCTCTCTCTC 86
CC
CC RESULT 334
CC ID ID14054 standard; cDNA; 2165 BP.
CC AC T14054_1996 (first entry)
CC CC 0-NON-ALB signal particle, SP254 gene cDNA clone
CC DN Signal recognition particle; recombinant; genetic engineering;
CC KE secretion; filamentous fungi; transformation; Aspergillus;
CC KW fungal enzymes; ds.
CC RW Best Local Similarity 100.0%; Pred. No. 3,47e+02;
CC FR Key Aspergillus also Location/Qualifiers
CC FD cds 86..1693
CC FT /*tag= a
CC FT WNG6106315-6.
CC FD 13-APR-1996.
CC PP 02-OCT-1995; U12628.
CC PR 03-OCT-1994; U12140.
CC PT (NOMO ) GENO MONTEK BIOTEC INC.
CC P1 WPI: 96-200357/2162 BP.
CC P2 WPI: 96-200357/2162 BP.
CC DR P-PDB: R3130.
CC PR Nucleic acid construct encoding elements of signal recognition particle
CC in filamentous fungi - used to increase prodn. of heterologous proteins
CC in fungi.
CC Claim 8: Ptg 3: 39pp: English.
CC T14053 and T14054 both encode an Aspergillus niger signal recognition
CC particle element, strictly derived from a genomic Bcl. clone and a cDNA
CC clone. The signal recognition particle (SRP) gene encoding
CC SP254 may be used for the recombinant production of a protein of
CC interest, in particular for the production of fungal enzymes e.g.
CC catalase, lactase, cellulase, xylanase, glucoamylase, pectinase or
CC chitinase. The signal recognition particle (SRP) gene encoding
CC components in transformed cells, improving the production of secreted
CC proteins.
CC Sequence 2165 BP: 463 A: 617 C: 578 G: 506 T:
CC
CC Query Match
CC 1.33: Score 14; DB 39; Length 2165;
CC Best Local Similarity 100.0%; Pred. No. 3,47e+02;

```

Query Match 1.38; Score 14; DB 23; Length 2165;
Best Local Similarity 100.00; Pred. No. 3.47e+02;

```
Matches      14: Conservative      0: Mismatches      0: Indels      0: Gaps      0:
Db    665   cgtcacaaagcaggga 678
      |||||
Cp    804   CGTCACAAGCAGGA 791
```

RESULT	335
ID	N92402 standard; DNA; 2177 BP.
AC	N92402;
DT	26-Nov-1990 (first entry)
DE	B subunit of human factor XIII.
DE	B sub-unit; factor XIII; blood clot stabilisation; fibrin polymer;
NW	B sub-unit; factor XIII; blood clot stabilisation; fibrin polymer;
KW	Crosslink; aciderodema; haemorrhage; ulcerative colitis; ss.

OS	Homo sapiens.
NCBI	taxid:9606
DB	Cds
RefSeq	Location/Qualifiers
chr1	2..1521
Exon	/tag a
Intron	/product b subunit
mRNA	/tag b
cDNA	/label leader sequence
cds	59..1521
exon	/tag c
intron	/product mature b subunit

AU8778694-A.
 31-MAR-1988 076654
 11-SEP-1986 US-900512
 (SYNO) Zymogenetics Inc.
 David E Seale ML, Ichinose A, Holly JA, Parker CE;
 Weir BG-140537/21.
 New DNA sequences encoding sub-units of factor 13- and corresponding
 expression vectors and transformed host cells.

p8 Disclosure: An English nucleotide sequence (1979-1981), is followed by a German (TAG) and a Russian (TAS) tail of 5 bp. The polyadenylation or processing signal of AATAA was identified 19 nucleotides upstream from the poly(A) tail. When cultured the host cells will produce the polypeptide which crosslinking fibrin polymers¹, factor III is useful therapeutically in cases of scleroderma, hemorrhage, ulcerative colitis etc., and can be prep'd. in large amounts without risk of viral contamination.

Query Match 1.3%; Score 14; DB 1; Length 2177;
Best Local Similarity 100.0%; Pred. NO. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
See also N24301 and N24303.
SQ Sequence 2177 BP; 758 A; 358 C; 434 G; 627 T;

Tue Nov 17 08:55:26 1998

```

Best Local Similarity 100.0%:      pred. NO. 3.47e+02;
Matches 14: Conservative          0; Mismatches 0; Indels 0; Gaps 0;

db 698 gtctgagaggaag 711
      |||.....|

```

Accession	Strain	Result
980 GUTTAGAGGAG 993	Q10958	standard; DNA; 2196 BP.
	Q10958: AC	
	Q10958: AC	
	23-MAR-1991	(first entry)
	Gen encoding acylamine acid-isolating enzyme-like polypeptide	

AKW	MARE; pig liver; ss.	
OS	Sus scrofa.	
PH	Key	Location/Qualifiers
FF	1. 2196	
FF	cda	
FF		/*tag= a
FF		/product= MARE-like polypeptide
PD	J03030673-A.	
PD	08-FEB-1991.	

29-JUN-1989: 165216.
PFF
29-JUN-1989: 04:165216
PR
29-JUN-1989: 04:165216
R
WPI: 91 084340/12
P-PDS: RLC056.
AC

Amylase acid-insoluble enzyme-like polypeptide - prepd. by
T. Kato et al., National Institute of Advanced Industrial Science
and Technology, Tsukuba, Japan.

Clsim 1, Pdg 2, 10pp; Japanese.
PS
The sequence was obtd. from six clones isolated from a pig liver
cDNA library. The clones, lambda AAHE 419, 450, 451, 452 and
CCC

Seq	12Z COVER	entire sequence of	the gene	677 G;	487 T;
	Sequence	2198 bp;	438 A;	594 C;	
	Query Match	Score 140	Score 140	DB 2	Length 2196;
	Best Local Similarity	100.0%	100.0%	Prod No. 3.47e+02;	
	Matches	44; Conservative	0; Mismatches	0; Indels	0; Gaps
bb	1216 tcaggtgaagctgt 1229				
ay	596 tcaagctgaagctg 609				

RESULT	338
ID	Q49052 standard; DNA; 2198 BP.
CD	35902-1994 (first entry)
CC	35902-1994 (first entry)
CE	Granulin coding sequence
CF	Granulin; keratinocytes
CG	Granulin; keratinocytes; wound healing; inhibition; peptide;
CH	granulocytes; leucocytes; ss.
CI	Homo sapiens.
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CZ	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	</

Page 335

```

Best Local Similarity 100.0%:      pred. NO. 3.47e+02;
Matches 14: Conservative          0; Mismatches 0; Indels 0; Gaps 0;

db 698 gtctgagaggaag 711
      |||.....|

```

Accession	Strain	Result
980 GUTTAGAGGAG 993	Q10958	standard; DNA; 2196 BP.
	Q10958: AC	
	Q10958: AC	
	23-MAR-1991	(first entry)
	Gen encoding acylamine acid-isolating enzyme-like polypeptide	

AKW	MARE; pig liver; ss.	
OS	Sus scrofa.	
PH	Key	Location/Qualifiers
FF	1. 2196	
FF	cda	
FF		/*tag= a
FF		/product= MARE-like polypeptide
PD	J03030673-A.	
PD	08-FEB-1991.	

29-JUN-1989: 165216.
PFF
29-JUN-1989: 04:165216
PR
29-JUN-1989: 04:165216
R
WPI: 91 084340/12
P-PDS: RLC056.
AC

Amylase acid-insoluble enzyme-like polypeptide - prepd. by
T. Kato et al., National Institute of Advanced Industrial Science
and Technology, Tsukuba, Japan.

Clsim 1, Pdg 2, 10pp; Japanese.
PS
The sequence was obtd. from six clones isolated from a pig liver
cDNA library. The clones, lambda AAHE 419, 450, 451, 452 and
CCD

Seq	12Z COVER	entire sequence of	the gene	677 G;	487 T;
	Sequence	2198 bp;	438 A;	594 C;	
	Query Match	Score 140	Score 140	DB 2	Length 2196;
	Best Local Similarity	100.0%	100.0%	Prod No. 3.47e+02;	
	Matches	44; Conservative	0; Mismatches	0; Indels	0; Gaps
bb	1216 tcaggtgaagctgt 1229				
ay	596 tcaagctgaagctg 609				

RESULT	338
ID	Q49052 standard; DNA; 2198 BP.
CD	35902-1994 (first entry)
CC	35902-1994 (first entry)
CE	Granulin coding sequence
CF	Granulin; keratinocytes
CG	Granulin; keratinocytes; wound healing; inhibition; peptide;
CH	granulocytes; leucocytes; ss.
CI	Homo sapiens.
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CZ	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	
DK	
DL	
DM	
DN	
DO	
DP	
DQ	
DR	
DS	
DT	
DU	
DV	
DW	
DX	
DY	
DA	
DB	
DC	
DD	
DE	
DF	
DG	
DH	
DI	
DJ	</

Tue Nov 17 08:55:26 1998

```

PH Key Location/Qualifiers
PT misc_difference 1453..1455
PT /tag=a
PT
PT
PT
PT -tag Alignment 117..119
PT

```

```

P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16
P17
P18
P19
P20
P21
P22
P23
P24
P25
P26
P27
P28
P29
P30
P31
P32
P33
P34
P35
P36
P37
P38
P39
P40
P41
P42
P43
P44
P45
P46
P47
P48
P49
P50
P51
P52
P53
P54
P55
P56
P57
P58
P59
P60
P61
P62
P63
P64
P65
P66
P67
P68
P69
P70
P71
P72
P73
P74
P75
P76
P77
P78
P79
P80
P81
P82
P83
P84
P85
P86
P87
P88
P89
P90
P91
P92
P93
P94
P95
P96
P97
P98
P99
P100
P101
P102
P103
P104
P105
P106
P107
P108
P109
P110
P111
P112
P113
P114
P115
P116
P117
P118
P119
P120
P121
P122
P123
P124
P125
P126
P127
P128
P129
P130
P131
P132
P133
P134
P135
P136
P137
P138
P139
P140
P141
P142
P143
P144
P145
P146
P147
P148
P149
P150
P151
P152
P153
P154
P155
P156
P157
P158
P159
P160
P161
P162
P163
P164
P165
P166
P167
P168
P169
P170
P171
P172
P173
P174
P175
P176
P177
P178
P179
P180
P181
P182
P183
P184
P185
P186
P187
P188
P189
P190
P191
P192
P193
P194
P195
P196
P197
P198
P199
P200
P201
P202
P203
P204
P205
P206
P207
P208
P209
P210
P211
P212
P213
P214
P215
P216
P217
P218
P219
P220
P221
P222
P223
P224
P225
P226
P227
P228
P229
P230
P231
P232
P233
P234
P235
P236
P237
P238
P239
P240
P241
P242
P243
P244
P245
P246
P247
P248
P249
P250
P251
P252
P253
P254
P255
P256
P257
P258
P259
P260
P261
P262
P263
P264
P265
P266
P267
P268
P269
P270
P271
P272
P273
P274
P275
P276
P277
P278
P279
P280
P281
P282
P283
P284
P285
P286
P287
P288
P289
P290
P291
P292
P293
P294
P295
P296
P297
P298
P299
P300
P301
P302
P303
P304
P305
P306
P307
P308
P309
P310
P311
P312
P313
P314
P315
P316
P317
P318
P319
P320
P321
P322
P323
P324
P325
P326
P327
P328
P329
P330
P331
P332
P333
P334
P335
P336
P337
P338
P339
P340
P341
P342
P343
P344
P345
P346
P347
P348
P349
P350
P351
P352
P353
P354
P355
P356
P357
P358
P359
P360
P361
P362
P363
P364
P365
P366
P367
P368
P369
P370
P371
P372
P373
P374
P375
P376
P377
P378
P379
P380
P381
P382
P383
P384
P385
P386
P387
P388
P389
P390
P391
P392
P393
P394
P395
P396
P397
P398
P399
P400
P401
P402
P403
P404
P405
P406
P407
P408
P409
P410
P411
P412
P413
P414
P415
P416
P417
P418
P419
P420
P421
P422
P423
P424
P425
P426
P427
P428
P429
P430
P431
P432
P433
P434
P435
P436
P437
P438
P439
P440
P441
P442
P443
P444
P445
P446
P447
P448
P449
P450
P451
P452
P453
P454
P455
P456
P457
P458
P459
P460
P461
P462
P463
P464
P465
P466
P467
P468
P469
P470
P471
P472
P473
P474
P475
P476
P477
P478
P479
P480
P481
P482
P483
P484
P485
P486
P487
P488
P489
P490
P491
P492
P493
P494
P495
P496
P497
P498
P499
P500
P501
P502
P503
P504
P505
P506
P507
P508
P509
P510
P511
P512
P513
P514
P515
P516
P517
P518
P519
P520
P521
P522
P523
P524
P525
P526
P527
P528
P529
P530
P531
P532
P533
P534
P535
P536
P537
P538
P539
P540
P541
P542
P543
P544
P545
P546
P547
P548
P549
P550
P551
P552
P553
P554
P555
P556
P557
P558
P559
P560
P561
P562
P563
P564
P565
P566
P567
P568
P569
P570
P571
P572
P573
P574
P575
P576
P577
P578
P579
P580
P581
P582
P583
P584
P585
P586
P587
P588
P589
P590
P591
P592
P593
P594
P595
P596
P597
P598
P599
P600
P601
P602
P603
P604
P605
P606
P607
P608
P609
P610
P611
P612
P613
P614
P615
P616
P617
P618
P619
P620
P621
P622
P623
P624
P625
P626
P627
P628
P629
P630
P631
P632
P633
P634
P635
P636
P637
P638
P639
P640
P641
P642
P643
P644
P645
P646
P647
P648
P649
P650
P651
P652
P653
P654
P655
P656
P657
P658
P659
P660
P661
P662
P663
P664
P665
P666
P667
P668
P669
P670
P671
P672
P673
P674
P675
P676
P677
P678
P679
P680
P681
P682
P683
P684
P685
P686
P687
P688
P689
P690
P691
P692
P693
P694
P695
P696
P697
P698
P699
P700
P701
P702
P703
P704
P705
P706
P707
P708
P709
P710
P711
P712
P713
P714
P715
P716
P717
P718
P719
P720
P721
P722
P723
P724
P725
P726
P727
P728
P729
P730
P731
P732
P733
P734
P735
P736
P737
P738
P739
P740
P741
P742
P743
P744
P745
P746
P747
P748
P749
P750
P751
P752
P753
P754
P755
P756
P757
P758
P759
P760
P761
P762
P763
P764
P765
P766
P767
P768
P769
P770
P771
P772
P773
P774
P775
P776
P777
P778
P779
P780
P781
P782
P783
P784
P785
P786
P787
P788
P789
P790
P791
P792
P793
P794
P795
P796
P797
P798
P799
P800
P801
P802
P803
P804
P805
P806
P807
P808
P809
P810
P811
P812
P813
P814
P815
P816
P817
P818
P819
P820
P821
P822
P823
P824
P825
P826
P827
P828
P829
P830
P831
P832
P833
P834
P835
P836
P837
P838
P839
P840
P
```

DR WPI: 93-720328/40.
DR p-5203: 948673.
PT New synthetic skin granulin peptide(s) from leucocyte(s) - are
PT keratinocyte inhibitors useful topically for wound healing
PS Disclosure: Figure 4c: 53pp; English.
CC The granulin inhibits keratinocytes and is useful in formulating
CC for promoting the healing of wounds.
SO Sequence: 2198 bp: 188 A: 708 C: 646 G: 446 T:

```

Query Match      1.30; Score 14; DB 9; Length 2198;
Best Local Similarity 100.00; Pred. 3.4e+00;
Matches 14; Conservative 0; Mismatches 0; Indels 0;

DB      828 cctgcactgcgtcc 841
      |||||
QY      891 CCTGCACTGCCTGCC 904

```

RESULTS	339
AC	standard; cDNA; 2218 BP.
AC	T62417
DT	02-JUL-1997 (first entry)
DE	S.-JUGLIA phenolpyruvate carboxylase coding sequence.
DE	SPERMATOPHYTES carboxylase; enzyme; photosynthesis; com-
KW	Poa annua; Poaceae; Poaceae; Poaceae; Poaceae; Poaceae;
KW	Eragrostis; buffalograss; Zea mays; Zea mays; Zea mays;
OS	Spain anglica.
PH	key
PH	Location/Qualifiers

FT	/product= phosphoenolpyruvate carboxylase
PT	/catalytic_activity= 2187..2192
FT	/polya_signal
PT	*tag= b
FT	J0905586-A.
PN	PN
PP	18-JUN-1996; 157060.
PR	22-JUN-1995; JP-156285.
PA	(NITA-) JAPAN TATU GURASU KK.

Page 336

PH		Location/Qualifiers
PT	1056-1055	
FT	/tag= a	
PT	/transl_except= ARG	encodes value.
PT	-1054-1053	

F7		misc_difference int: 1126
F8	/seqs &	
F9	/trans_except= CAG encodes Glycine.	
P1		
P2	W0315195-A.	
P3	05-APC-1997.	
P4	05-APC-1997.	
P5	28-FEB-1997; CA0089.	
P6	03-FEB-1997; IS-829213.	
P7	(SOLG) SOLG00H S.	
P8	SOLG00H S.	

DR WPI: 93-320328/40.
 DR P-9503: 846673.
 PT New cationic rich granulin peptide(s) from leucocytes(s) - are
 PT keratinocyte inhibitors useful topically for wound healing
 PS Disclosure: Figure 4c; 53pp; English.
 CC The granulin inhibits keratinocytes and is useful in formulations
 CC for promoting the healing of wounds.
 SO Sequence: 2198 BP: 398 A: 708 C: 646 G: 446 T:

```

Query Match      1.34; Score 14; DB 9; Length 2190;
Best Local Similarity 10.04; Pred. No. 3.47e+02;
Matches      14; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      828 cctgcactgcgc 841
        |||||
Qy      891 CCTGCACCTGCC 904

```

RESULTS 339
The standard deviation was 0.007. The mean value was 0.007. The standard error was 0.007.
AC 763417.
DT 02-JUL-1997 (first entry)
DE S. anglica phosphoenolpyruvate carboxylase coding sequence.
EN Spontaneous mutation in the gene encoding phosphoenolpyruvate carboxylase; condenses
KW panlograses; buftolograses; tsoisina japonica; es.
OS Spartina anglica.
Key Location/Qualifiers
PH

FT	polya_signal	/tag= a	/product= phosphoenolpyruvate carboxykinase
FT	2187..2192	/tag= b	
FT	309055866..A.		
PN	18-JUN-1996		
PP	18-JUN-1996		157060.
PR	22-JUN-1995		JP-156285.
PA	(NITA-) JAPAN TAFU GURASU KK.		

PN EP-58093-A.
 PD 20-AUG-1993: 306628
 PR 21-AUG-1993: JP-243984.
 PR (TAKI) TAKARA SHUZO CO LTD.
 PI Nishikawa A., Taniguchi N.
 DR P-PSDB: R48993/09.
 New gene for human glycosyltransferase GGT-III - and related
 PT vectors and transformed cells, useful in diagnosis of cancer
 PS Claim 1: Page 10: 14pp: English.
 CC Glycosyltransferase gene may be used for the expression of human
 CC GGT-III in cells. The gene encodes a protein of 378 amino acids.
 CC Glucosaminyltransferase in host cells. The gene and its
 CC decomposition products may be used in cancer diagnosis.
 CC Sequence 2247 BP: 368 A: 751 C; 750 G; 378 T;
 Query Match 1.38; Score 15; DB 10; Length 2247;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01; Mismatches 0; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 563 agacacgctgctt 577
 Cp 1015 AGACACGCTGCTGAG 1001
 RESULT 344
 ID T60336 standard; DNA: 2250 BP.
 AC T60336.
 DT 29-MAY-1997 (first entry)
 CC 0 mutant coding sequence.
 CC TIM-1 mutant coding sequence.
 CC sleep-wake cycle; light sensitive; environmental cycle; circadian rhythm;
 CC Drosophila period gene; transcription factor; PAS domain; depression;
 CC narcolepsy; PER; jet lag; MNP; therapy; as.
 CC P-PSDB: R48993/09.
 CC 26-SEP-1996.
 CC 20-MAR-1996; 003830.
 CC 10-MAR-1995; US-408519.
 CC 02-NOV-1995; US-552554.
 CC (UTPE-) UNIV PENNSYLVANIA.
 CC (UTRO-) UNIV ROCKEFELLER.
 CC WPI-96-443181C JL, Sehgal A, Voeshall LB, Young MW;
 CC P-PSDB: M14053.
 CC Nuclear translocation protein which binds to protein involved in
 CC circadian rhythms - used to develop prods. for diagnosis, prevention
 CC of sleep-wake cycle disorders associated with circadian rhythms, e.g.
 CC jetlag and narcolepsy

PS Example 9: 13pp: English.
 CC This sequence represents the coding sequence for the 01 mutant of the
 CC Drosophila melanogaster "timeless" (TIM) protein. The TIM01 mutant is an
 CC inversion of the TIM protein. The TIM protein is a nuclear translocation protein (NTP) of the
 CC protein involved in circadian rhythms, and cyclic transcription patterns
 CC related to the sleep-wake cycle. The BTP is also preferentially light
 CC sensitive, and has a stabilizing effect on the circadian rhythm protein.
 CC entrainment to environmental cycles of light. TIM has specific binding
 CC activity for the Drosophila period (PER) gene. PER is a nuclear protein
 CC which has homology to the family of transcription factors containing the
 CC PAS domain, and is a protein involved in circadian rhythms, but the with
 CC a circadian rhythm. The protein is also phosphorylated with a circadian
 CC rhythm. The BTPs of the invention, their fragments, agonists, mimics and
 CC antagonists are useful for preventing and/or treating disorders of a
 CC circadian rhythm such as depression, narcolepsy or jet lag. The products
 CC classify groups of individuals with the disorders.
 CC Sequence 2250 BP: 590 A: 616 C; 597 G; 446 T;
 Query Match 1.38; Score 14; DB 29; Length 2250;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 892 ttgagagctgac 905
 QY 84 TTCAGAGGCTGAG 97
 RESULT 345
 ID T71741 standard; cDNA to mRNA; 2251 BP.
 AC T71741.
 DT 09-OCT-1996 (first entry)
 CC Tilaapia prolactin receptor cDNA.
 CC Tilaapia prolactin (TILPRL) receptor; hormone; agonist; antagonist;
 CC rat prolactin receptor; as. Synchronization; teleost; bony fish.
 CC Cypriiniformes; as.
 CC Oretochromis niloticus.
 CC Key Key Location/Qualifiers
 CC cda /tag=099
 CC /tag=099
 CC /product= pre_PRL_receptor
 CC /tag= b
 CC /product= PRL_receptor
 CC FR2724181-A1.
 CC 08-MAR-1996.
 CC 01-SEP-1994; 010535.
 CC 01-SEP-1994; FR-010535.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PI Eery M, Kelly PA, Prunet P, Sandra O;
 DR WPI-96-15312A/18
 CC Fish prolactin receptor - useful for detecting prolactin agonists
 CC and antagonists.
 CC Claim 8: Page 21-22: 3pp: French.
 CC cDNA sequence coding for a fish prolactin (PRL) receptor was
 CC isolated from a fish brain and expressed in a mammalian cell line.
 CC expression library following screening with a PRL antibody, the PRLP.
 CC The cDNA insert was found to contain an open reading frame for a
 CC 630 amino acid protein. The mature protein (606 amino acids) has
 CC a hydrophobic signal sequence (1-24) and a cytoplasmic domain (25-606).
 CC eukaryotic cells. The mature protein (606 amino acids) has
 CC identifying agonists and antagonists of PRL which have potential
 CC applications in fish farming.
 CC Sequence 2251 BP: 610 A: 577 C; 573 G; 491 T;
 Query Match 1.38; Score 14; DB 21; Length 2251;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1450 agacacgcttctt 1463
 Cp 210 AGACCTGCTCTCT 197
 RESULT 346
 ID Q05673 standard; cDNA; 2252 BP.
 AC Q05673.
 DT 03-JAN-1991 (first entry)
 CC Desmodium rotundifolium.
 CC Thrombolytic agent: fibrinolysis; saline; vampire bat; as.
 CC Desmodium rotundifolium.
 CC Key Key Location/Qualifiers
 CC cda /tag=1527
 CC /tag=1527
 CC /product=v-PAL_alpha
 CC EP-383417-A.
 CC 22-AUG-1990. 260043
 CC 11-FEB-1989; DE-904580.
 CC 30-MAY-1989; DE-917949.
 CC (SCID) SCHERING AG.
 CC Balch B, Donner P, Schleuping WD, Alagon A, Boisdol W;
 CC WPI-90-255678/14.
 CC P-PSDB: R06457.
 CC New fibrin specific thrombolytic agent v-PAL - isolated from
 CC bat. Desmodium sp. bats.
 CC Claim 5: Fig 26a, 49pp; English.

CC The sequence was obtd. from the longest of the alpha 1 gp. of
 CC v-PAL clones isolated from a cDNA library prepd. from RNA from
 CC the salivary glands of vampire bats. The v-PAL protein produced
 CC on expression of the gene is a high molecular form consisting of
 CC a single chain of 666 amino acids. The protein has a signal domain and a protease
 CC domain. It is a fibrinolytic specific and dissolves blood clots.
 CC See also Q05674-Q05678.
 CC Sequence 2252 BP: 632 A: 548 G; 570 G; 502 T;
 Query Match 1.38; Score 14; DB 31; Length 2252;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 653 ctatctctctctct 666
 QY 687 CAAACCTCTCTCT 700
 RESULT 347
 ID Q12189 standard; DNA; 2252 BP.
 AC Q12189.
 DT 11-SEP-1991 (first entry)
 CC Rabbit sperm-specific nuclear auto-antigenic protein gene.
 CC Rabbit sperm-specific nuclear auto-antigenic protein gene.
 CC Oryctolagus cuniculus.
 CC Key Key Location/Qualifiers
 CC cda /tag=2138
 CC /tag=2138
 CC /product= rabbit SNP
 CC W09108289-A.
 CC 13-JUN-1991.
 CC 21-NOV-1990; U06863.
 CC 14-NOV-1990; US-41036.
 CC (UNIC-) UNIV NORTH CAROLINA.
 CC Grand MG, Kimmelman LJ, Welch JE;
 CC WPI-91-193195/26.
 CC P-PSDB: R12540.
 CC Rabbit sperm-specific nuclear auto-antigenic protein gene - encoding sperm-specific nuclear
 CC auto-antigenic protein used as immuno-contraceptive and to
 CC diagnose auto-immune infertility
 CC Claim 3; Fig 1; 46pp: English.
 CC This sequence was isolated by screening rabbit seminiferous tubule
 CC cDNA libraries with a rabbit sperm-specific nuclear auto-antigenic protein gene (R1.2)
 CC identified by screening for rabbit sperm membrane autoantigen.
 CC See also Q12190 and Q12191.
 CC Sequence 2252 BP: 764 A: 376 G; 645 G; 467 T;
 Query Match 1.38; Score 14; DB 2; Length 2252;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      1010 agagctgacgcagca 1023
QY      88 AGAGCTCAGGCAG 101

RESULT 348
ID       Q56758 standard; DNA: 2755 BP.
AC       Q56758.
DE       C2-PCR-1994 (first entry) al.
DI       Chitinase derivative: beta-1,4-glycoside bond; chitin;
KW       Km Beta-N-acetylhexosaminidase activity; yeast; asn.
OS       Rhizopus niveus
SS       Location/Qualifiers
FT       cds              138..1750
FT       CDS               /tag= a
FT       FT                /product= Chitinase
FT       FT                /tag= b
FT       exon             243..308
FT       FT               /tag= c
FT       FT               /tag= d
FT       exon             306..499
FT       FT               /tag= e
FT       intron           499..566
FT       FT               /tag= e
FT       exon             567..1750
FT       FT               /tag= f
PN       J06046849-A.
DP       22-FEB-1994.
PF       28-JUL-1992. 201427.
PR       28-JUL-1992. 201427.
PS       WPI: 94-09015/12
DR       P-PSDB: R48669.
DT       New chitinase enzyme and coding sequence - cleaves beta-1,4-glycoside
TS       Claim 11, Page 21,22: 68PP: Japanese Patent Application No. 557156-61 encode chitinase derivatives which
CC       cleave the beta-1,4-glycoside bond of chitin but have substantially
CC       no beta-N-acetylhexosaminidase activity. These sequences may be
CC       used to identify and culture for the production of the
CC       chitinase products.
CC       Chitinase 2255 BP: 61A C: 448 C: 446 G: 747 T:
SQ
Query Match 138 Score 14; DB 10; Length 2255;
Best Local Similarity 100.0%; Pred No. 3,47+02; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;
Db      1283 tccatcaccgtcgatt 1296
QY      470 TCATCACCCTGATT 1283

```

```

Tue Nov 17 08:55:26 1998

US-08-887-977-9.1mg

KW Kaposi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;
KS therapy; diagnosis; antisense; triplex; vaccine; AIDS; as.
OS Kapoai's sarcoma associated herpesvirus.
PN W06137-1991.
PP 21-NOV-1995: U15138.
PR 21-NOV-1994: US-343101.
PR 11-APR-1995: US-420235.
PA (C)CO 1 UNIV COLOMBIA NEW YORK.
PI P. A.
WP1: 95-268320/27.
DA WPI: 95-268320/27.
DR P-PSDB: R97834.
DT Herpes virus associated with Kaposi's sarcoma - also definitive DNA
PT sequences, useful for diagnosis of and to develop prods. for
TS 1995.
P6 Claim 17, Page 407-208; 277pp; English.
P8 Lambda clone KS5 (T30681) has 17 protein-coding regions (T30682-98),
CC 15 of which are complete ORFs, including ORF24 (T30686) (aka 5500-
CC 7758), complementary strand). KS5 is a fragment of a newly identified
CC ORF8 were identified by sequence and positional homology to know
CC herpesvirus sequences. Probes that bind specifically to the KS
CC associated herpesvirus can be used for KS diagnosis. Antisense or
CC triplex oligonucleotides are useful for prophylaxis or treatment of
CC KS.
CC Reading frames are useful as vaccines.
SQ Sequence 2259 BP; 586 A; 632 C; 551 G; 490 T;
Query Match 138; Score 15; DB 21; Length 2259;
Matched Similarity 100%;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 676 acetatagcaggcttc 690
Cc 617 AACATGACAGCTTC 603

RESULT 351 standard: DNA: 2260 BP.
ID T27731
AD T27731
DT 08-NOV-1996 (first entry)
DE Mutant insulin-stimulated protein kinase (ISPK-1) coding sequence.
DE Insulin stimulated protein kinase; ISPK: detection; screening;
DE diagnosis; resistance; diabetes mellitus; MIMD; as.
OS Key antigens. Location/Qualifiers
FH Key antigens. Location/Qualifiers
FF cds
FT 1..2232 a
FT /product= Insulin-stimulated protein Kinase.
PD W0615233-AL
PD 23-MAY-1998.
```

```

RESULT 349
ID T16816 standard; DNA; 2259 BP.
AC T16816; 1996 (first entry)
DZ Kaposi's sarcoma associated herpesvirus ORP24.
KW Kaposi's sarcoma; KSRY; lymphoma; AIDS; vaccine; diagnosis; therapy;
# Kaposi's sarcoma associated herpesvirus.
PD WP066159A.
PS 23-FEB-1996.
PF 11-AUG-1985; D10194.
PP 18-AUG-1994; US-292165.
PR 18-AUG-1994; US-292165.
PR 11-APR-1995; US-420235.
PA (UTCC) UNIV COLOMBIA NEW YORK.
FA (GRAM) GRANT D E.
FA (GRAM) YANKEE J.
PI Chang Y, MOORE PS, Grant DE, Viale L.
PFI WPI: 96-151362/15.
P PDSB: R39610.
DR Herpesvirus DNA associated with Kaposi's sarcoma - also associated
DR with hairy-cell leukemia and vaccination.
PS Claim 17; Page 220-222; 305pp; English.
CC Abpd. from a KS lesion genomic library, includes 15 complete ORFs and
CC encodes at least 12 proteins; nucleocapsid protein p160, tegument
CC asimilar positional homologues ORP24 is akin to T16816 tat protein
CC encoded protein in KSHV10. KSHV DNA may be incorporated into a
CC vector and expressed in host cells to produce peptides useful in
CC immunoscreening assays and antibodies. The DNA may itself be used to
CC transfect suitable host cells and generate recombinant plasmids and
CC triplex molecules or diagnostic probes, or to raise transgenic
CC animals.
SQ Sequence 2259 BP; 586 A; 632 C; 551 G; 490 T;
Query Match 1.3%; Score 15; DB 20; Length 2259;
Best Local Similarity 100.0%; Fwd M. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Db 676 sacatccagacgaccttc 690
|||||
CP 617 AACATCAGCGCTTC 603

RESULT 350
ID T30686 standard; DNA; 2259 BP.
AC T30686; 1996 (first entry)
DZ Kaposi's sarcoma associated herpesvirus ORP24.
KW Kaposi's sarcoma; KSRY; lymphoma; AIDS; vaccine; diagnosis; therapy;
# Kaposi's sarcoma associated herpesvirus.
PD WP066159A.
PS 23-FEB-1996.
PF 11-AUG-1985; D10194.
PP 18-AUG-1994; US-292165.
PR 18-AUG-1994; US-292165.
PR 11-APR-1995; US-420235.
PA (UTCC) UNIV COLOMBIA NEW YORK.
FA (GRAM) GRANT D E.
FA (GRAM) YANKEE J.
PI Chang Y, MOORE PS, Grant DE, Viale L.
PFI WPI: 96-151362/15.
P PDSB: R39610.
DR Herpesvirus DNA associated with Kaposi's sarcoma - also associated
DR with hairy-cell leukemia and vaccination.
PS Claim 17; Page 220-222; 305pp; English.
CC Abpd. from a KS lesion genomic library, includes 15 complete ORFs and
CC encodes at least 12 proteins; nucleocapsid protein p160, tegument
CC asimilar positional homologues ORP24 is akin to T16816 tat protein
CC encoded protein in KSHV10. KSHV DNA may be incorporated into a
CC vector and expressed in host cells to produce peptides useful in
CC immunoscreening assays and antibodies. The DNA may itself be used to
CC transfect suitable host cells and generate recombinant plasmids and
CC triplex molecules or diagnostic probes, or to raise transgenic
CC animals.
SQ Sequence 2259 BP; 586 A; 632 C; 551 G; 490 T;
Query Match 1.3%; Score 15; DB 20; Length 2259;
Best Local Similarity 100.0%; Fwd M. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Db 676 sacatccagacgaccttc 690
|||||
CP 617 AACATCAGCGCTTC 603

```

```

US-08-087-977-9.rng

The Nov 17 08:55:26 1998

PF 10-NOV-1995; DK0448
PA 10-NOV-1994; DK-001294.
PR (NOVO ) NOVO-NORDISK AS.
PI Bjorbaek C, Schwelb SM, Pedersen O;
DR P-DBS; R35917/26.
PT Mutant DNA encoding insulin-stimulated protein kinase - useful to
PT diagnose pre-disposition to insulin resistance and possibly NIDDM
PS Clin 3; Page 30; 42pp; English.
CC CC can be used as a mutant insulin-stimulated protein kinase (ISPK-1)
CC genes. The method of detection comprises analysing an isolated DNA
CC by digestion with a restriction endonuclease which cleaves at the
CC site of the mutation. The obtained restriction pattern is compared
CC to a negative control restriction pattern (obtained from a wild type
CC 1998-10-01) and the difference between the two patterns is used to
CC from a mutant ISPK-1 gene. This method may be used to diagnose
CC pre-disposition to insulin resistance and possibly non-insulin-
CC dependent diabetes mellitus. This mutant sequence contains a T to
CC C substitution at the second position of codon 38 resulting in a
CC 1385 substitution in the mutated DNA. 417 C; 511 G; 628 T;
SQ Sequence 2260 BP; 704 A; 422 C; 417 G; 628 T;

Query Match 1.34; Score 14; DB 23;
Accession Local Uniquity 100.0%; Prod No: 3.47e-02; Length 2260;
Matrix 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1346 cagtagcagatt 1359
QY 14 CAGTAGCAGATT 27

RESULT 352
ID T27330 standard; DNA; 2260 BP.
CC 1346 cagtagcagatt 1359
DB 08-NOV-1996 (first entry)
DE Insulin-stimulated protein kinase (ISPK-1) coding sequence.
DE Insulin-stimulated protein kinase; ISPK; detection; screening;
DE diagnosis; resistance; diabetes mellitus; NIDDM; ds.
KW Key words.
KW Key words.
FT Key Location/Qualifiers
FT cds 1..2223
FT /tag= a
FT /product= Insulin-stimulated protein kinase.
FT WP0415233.A1.
FT 23-MAY-1996
FT 23-MAY-1996
PP 10-NOV-1995; DK0448.
PR 10-NOV-1994; DK-001294.
PA (NOVO ) NOVO-NORDISK AS.
PI Bjorbaek C, Schwelb SM, Pedersen O;
DR P-DBS; R35917/26.
PT Mutant DNA encoding insulin-stimulated protein kinase - useful to
PT diagnose pre-disposition to insulin resistance and possibly NIDDM
PS Clin 3; Page 30; 42pp; English.
CC CC can be used as a mutant insulin-stimulated protein kinase (ISPK-1)
CC genes. The method of detection comprises analysing an isolated DNA
CC by digestion with a restriction endonuclease which cleaves at the
CC site of the mutation. The obtained restriction pattern is compared
CC to a negative control restriction pattern (obtained from a wild type
CC 1998-10-01) and the difference between the two patterns is used to
CC from a mutant ISPK-1 gene. This method may be used to diagnose
CC pre-disposition to insulin resistance and possibly non-insulin-
CC dependent diabetes mellitus. This mutant sequence contains a T to
CC C substitution at the second position of codon 38 resulting in a
CC 1385 substitution in the mutated DNA. 417 C; 511 G; 628 T;
SQ Sequence 2260 BP; 704 A; 422 C; 417 G; 628 T;

```


CC		that exhibit activity against a broad range of intestinal and
CC		opportunistic pathogens Rat cryptidin-1, -2, and -3 genes (see
CC		also T30736 and T30738) were isolated from a Sprague-Dawley
CC		genomic library cloned in EMBl3. The cryptidin-2 gene, or
CC		cryptidin-2, was used as a probe to identify clones that produce
CC		large amounts of cryptidin-2 for use in treating inflammatory
CC		pathologies of the intestine.
CC		Sequence 2408 BP; 640 A; 531 C; 557 G; 680 T;
SQ		Query Match 1.3%; Score 15; DB 21; Length 2408;
BEST		Local Similarity 100.0%; Pred.No. 8.37e+01;
MATCHES	15:	Conservative O; Mismatches O; Indels O; Gaps O;
Dd	1016	gacatcnaaggat 1030
Cp	660	gacatcnaaggat 646
RESULT	369	ID Q69033 standard: cDNA: 2409 BP.
AC	Q69033:	1995 (x1 ref entry)
DF	MA-MAR-1995	(x1 ref entry)
KW	Porcine:	TRK; cell surface; receptor; tyrosine kinase; transmembrane;
KW	Trk:	murine; expression; vector; pMX-neo; oncogene; proto-oncogene;
KW	oncogenesis;	antitumor agent; ds.
KW	Key words:	Location/Qualifiers
FT	cda	1..2187
FT	/tag:	a
FT	/product:	murine TrkC protein
FT	/tag:	b
FT	misc_feature	1030..1075
FT	/note:	"A region of 75 bases found in variant TrkC K3
FT	/note:	gene replacing bases 1801-42 of the murine
FT	/note:	sequence. This region contains 9 codons where
FT	/note:	the k3 sequence differs from the protein here
FT	/note:	TrkC K2 protein have been replaced by the deduced
FT	/note:	a.a.601-75 of sequence shown in #741618."
FT	US534895A:	
PP	DN	28-SEP-1994.
PP	PP	08-JUL-1991: 726456.
PP	PP	08-JUL-1991: US-726468.
PR	PR	07-JUL-1992: US-912957.
PR	PR	07-JUL-1992: US-912957.
DR	DR	Barbacid WJ, Lamballe P,
DR	DR	WPI; 94-302202/37.
DR	DR	P-PDSB: R741618.
PT	PT	The nucleated acid encoding TrkC protein and related vectors - and
PT	PT	identifying anticancer agents
PT	PT	studying onco-genetic and

Tue Nov 17 08:55:26 1998

[illegible]

Tue Nov 17 08:55:26 1998

US-08-087-977-9.1pg

[illegible]

115-08-887-977-9 vno

	caps. for optoid (antagonist) cDNA	1.3% Beat Local similarity 100.0%; Pred. No. 3.47e+02; Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Pr	Pf-198P-E6 Engineered clones, 100,000 human genomic clones and a similar number of mouse genomic clones were probed with the 1.1 kb mouse delta opioid receptor clone DOR-1 Pst/XbaI fragment.	
Cc	CC One mouse clone and three human genomic clones were screened by Southern blotting techniques from which one was identified indicated that there different genes were represented by the human genomic clones which were designated H3, H14 and H20. H14 maps to chromosome B. It encodes the human kappa opioid receptor.	
SQ	Sequence 2447 bp 683 A: 512 C: 458 G:	Score 14; DB 10; Length 2447; Pred. No. 3.47e+02; Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Dd	Db 2259 ataaagaaccca 2112	
Oo	Op ATTTTTTTTTTTT ATATTGAGGCCA 634	
Rt	RESULT 372	
ID	ID Q03453 standard; cDNA: 2448 DP.	
DE	De 20-JUL-1990 (first entry)	
DD	DD cDNA molecule which hybridises to the cDNA of murine IRF-1 encoding a protein with IFN-1 activity.	
DB	DB Interferon beta gene; cis-elements: se.	
FH	FH Key Location/Qualifiers	
Ft	Ft cds 154..1200	
Ft	/seq= a	
Ft	/product=IRF-1	
Fm	Fm EP-35202-A.	
Fm	Fm 28-FEB-1990.	
PP	PP 24-MOV-1988; 115602.	
PA	PA GenBank accession #14602. (TMV)/ Tsai-guchi T.	
Tg	Tg Taniguchi T.	
PR	Pf-P1: 90-060144/O9.	
DR	P-PEBIS: R03589	
DR	DR The above sequence was induced activation of interferon-beta gene transcription by interacting with cis-elements.	
Pf	Pf Claim 11; Page 41-45; 65pp; English.	
Ps	PS The recombinant virus produced induction of interferon-beta gene transcription by interaction with cis-elements.	
OC	OC Of recombination molecules are used for production of ISF-1.	
CC	CC See also O03452, -51 and -55; and PF-355190-A.	
Se	Sequence 2448 BP; 692 A; 630 C; 510 G; 615 T; 1 Others;	

[illegible][illegible]

```

RESULT 375
ID ID Q14425 standard; cDNA; 2403 BP.
AC AC Q14425.1992 (first entry)
CD CD T3-DM-1392 (T3 DM-1392)
DE DE Tyrosine protein kinase; ss
KW KW Mus musculus.
NM NM
OR OR
PE PE
PP PP
PR PR
PS PS
PT PT
SC SC
SD SD
SE SE
SI SI
SL SL
SM SM
SN SN
SO SO
SR SR
SS SS
ST ST
SU SU
SV SV
SW SW
SX SX
SY SY
SZ SZ
TA TA
TB TB
TC TC
TD TD
TE TE
TF TF
TG TG
TH TH
TI TI
TJ TJ
TK TK
TL TL
TM TM
TN TN
TO TO
TP TP
TQ TQ
TR TR
TS TS
TT TT
TV TV
TW TW
TX TX
TY TY
TZ TZ
AA AA
AB AB
AC AC
AD AD
AE AE
AF AF
AG AG
AH AH
AI AI
AJ AJ
AK AK
AL AL
AM AM
AN AN
AO AO
AP AP
AQ AQ
AR AR
AS AS
AT AT
AU AU
AV AV
AW AW
AX AX
AY AY
AZ AZ
BA BA
BB BB
BC BC
BD BD
BE BE
BF BF
BG BG
BH BH
BI BI
BJ BJ
BK BK
BL BL
BM BM
BN BN
BO BO
BP BP
BQ BQ
BR BR
BS BS
BT BT
BU BU
BV BV
BW BW
BX BX
BY BY
BZ BZ
CA CA
CB CB
CC CC
CD CC
CE CE
CF CF
CG CG
CH CH
CI CI
CJ CJ
CK CK
CL CL
CM CM
CN CN
CO CO
CP CP
CQ CQ
CR CR
CS CS
CT CT
CU CU
CV CV
CW CW
CX CX
CY CY
CZ CZ
DA DA
DB DB
DC DC
DD DD
DE DE
DF DF
DG DG
DH DH
DI DI
DJ DJ
DK DK
DL DL
DM DM
DN DN
DO DO
DP DP
DQ DQ
DR DR
DS DS
DT DT
DU DU
DV DV
DW DW
DX DX
DY DY
DZ DZ
EA EA
EB EB
EC EC
ED ED
EE EE
EF EF
EG EG
EH EH
EI EI
EJ EJ
EK EK
EL EL
EM EM
EN EN
EO EO
EP EP
EQ EQ
ER ER
ES ES
ET ET
EU EU
EV EV
EW EW
EX EX
EY EY
EZ EZ
FA FA
FB FB
FC FC
FD FD
FE FE
FF FF
FG FG
FH FH
FI FI
FJ FJ
FK FK
FL FL
FM FM
FN FN
FO FO
FP FP
FQ FQ
FR FR
FS FS
FT FT
FU FU
FV FV
FW FW
FX FX
FY FY
FZ FZ
GA GA
GB GB
GC GC
GD GD
GE GE
GF GF
GG GG
GH GH
GI GI
GJ GJ
GK GK
GL GL
GM GM
GN GN
GO GO
GP GP
GQ GQ
GR GR
GS GS
GT GT
GU GU
GV GV
GW GW
GX GX
GY GY
GZ GZ
HA HA
HB HB
HC HC
HD HD
HE HE
HF HF
HG HG
HH HH
HI HI
HJ HJ
HK HK
HL HL
HM HM
HN HN
HO HO
HP HP
HQ HQ
HR HR
HS HS
HT HT
HU HU
HV HV
HW HW
HX HX
HY HY
HZ HZ
IA IA
IB IB
IC IC
ID ID
IE IE
IF IF
IG IG
IH IH
II II
IJ IJ
IK IK
IL IL
IM IM
IN IN
IO IO
IP IP
IQ IQ
IR IR
IS IS
IT IT
IU IU
IV IV
IW IW
IX IX
IY IY
IZ IZ
JA JA
JB JB
JC JC
JD JD
JE JE
JF JF
JG JG
JH JH
JI JI
JJ JJ
JK JK
JL JL
JM JM
JN JN
JO JO
JP JP
JQ JQ
JR JR
JS JS
JT JT
JU JU
JV JV
JW JW
JX JX
JY JY
JZ JZ
KA KA
KB KB
KC KC
KD KD
KE KE
KF KF
KG KG
KH KH
KI KI
KJ KJ
KK KK
KL KL
KM KM
KN KN
KO KO
KP KP
KQ KQ
KR KR
KS KS
KT KT
KU KU
KV KV
KW KW
KX KX
KY KY
KZ KZ
LA LA
LB LB
LC LC
LD LD
LE LE
LF LF
LG LG
LH LH
LI LI
LJ LJ
LK LK
LL LL
LM LM
LN LN
LO LO
LP LP
LQ LQ
LR LR
LS LS
LT LT
LU LU
LV LV
LW LW
LX LX
LY LY
LZ LZ
MA MA
MB MB
MC MC
MD MD
ME ME
MF MF
MG MG
MH MH
MI MI
MJ MJ
MK MK
ML ML
MM MM
MN MN
MO MO
MP MP
MQ MQ
MR MR
MS MS
MT MT
MU MU
MV MV
MW MW
MX MX
MY MY
MZ MZ
NA NA
NB NB
NC NC
ND ND
NE NE
NF NF
NG NG
NH NH
NI NI
NJ NJ
NK NK
NL NL
NM NM
NN NN
NO NO
NP NP
NQ NQ
NR NR
NS NS
NT NT
NU NU
NV NV
NW NW
NX NX
NY NY
NZ NZ
OA OA
OB OB
OC OC
OD OD
OE OE
OF OF
OG OG
OH OH
OI OI
OJ OJ
OK OK
OL OL
OM OM
ON ON
OO OO
OP OP
OQ OQ
OR OR
OS OS
OT OT
OU OU
OV OV
OW OW
OX OX
OY OY
OZ OZ
PA PA
PB PB
PC PC
PD PD
PE PE
PF PF
PG PG
PH PH
PI PI
PJ PJ
PK PK
PL PL
PM PM
PN PN
PO PO
PP PP
PQ PQ
PR PR
PS PS
PT PT
PU PU
PV PV
PW PW
PX PX
PY PY
PZ PZ
QA QA
QB QB
QC QC
QD QD
QE QE
QF QF
QG QG
QH QH
QI QI
QJ QJ
QK QK
QL QL
QM QM
QN QN
QO QO
QP QP
QQ QQ
QR QR
QS QS
QT QT
QU QU
QV QV
QW QW
QX QX
QY QY
QZ QZ
RA RA
RB RB
RC RC
RD RD
RE RE
RF RF
RG RG
RH RH
RI RI
RJ RJ
RK RK
RL RL
RM RM
RN RN
RO RO
RP RP
RQ RQ
RR RR
RS RS
RT RT
RU RU
RV RV
RW RW
RX RX
RY RY
RZ RZ
SA SA
SB SB
SC SC
SD SD
SE SE
SF SF
SG SG
SH SH
SI SI
SJ SJ
SK SK
SL SL
SM SM
SN SN
SO SO
SP SP
SQ SQ
SR SR
SS SS
ST ST
SU SU
SV SV
SW SW
SX SX
SY SY
SZ SZ
TA TA
TB TB
TC TC
TD TD
TE TE
TF TF
TG TG
TH TH
TI TI
TJ TJ
TK TK
TL TL
TM TM
TN TN
TO TO
TP TP
TQ TQ
TR TR
TS TS
TT TT
TV TV
TW TW
TX TX
TY TY
TZ TZ
UA UA
UB UB
UC UC
UD UD
UE UE
UF UF
UG UG
UH UH
UI UI
UJ UJ
UK UK
UL UL
UM UM
UN UN
UO UO
UP UP
UQ UQ
UR UR
US US
UT UT
UU UU
UV UV
UW UW
UX UX
UY UY
UZ UZ
VA VA
VB VB
VC VC
VD VD
VE VE
VF VF
VG VG
VH VH
VI VI
VJ VJ
VK VK
VL VL
VM VM
VN VN
VO VO
VP VP
VQ VQ
VR VR
VS VS
VT VT
VU VU
VV VV
VV VV
VW VW
VX VX
VY VY
VZ VZ
WA WA
WB WB
WC WC
WD WD
WE WE
WF WF
WG WG
WH WH
WI WI
WJ WJ
WK WK
WL WL
WM WM
WN WN
WO WO
WP WP
WQ WQ
WR WR
WS WS
WT WT
WU WU
WV WV
WW WW
WX WX
WY WY
WZ WZ
XA XA
XB XB
XC XC
XD XD
XE XE
XF XF
XG XG
XH XH
XI XI
XJ XJ
XK XK
XL XL
XM XM
XN XN
XO XO
XP XP
XQ XQ
XR XR
XS XS
XT XT
XU XU
XV XV
XW XW
XX XX
XY XY
XZ XZ
YA YA
YB YB
YC YC
YD YD
YE YE
YF YF
YG YG
YH YH
YI YI
YJ YJ
YK YK
YL YL
YM YM
YN YN
YO YO
YP YP
YQ YQ
YR YR
YS YS
YT YT
YU YU
YV YV
YW YW
YX YX
YY YY
YZ YZ
ZA ZA
ZB ZB
ZC ZC
ZD ZD
ZE ZE
ZF ZF
ZG ZG
ZH ZH
ZI ZI
ZJ ZJ
ZK ZK
ZL ZL
ZM ZM
ZN ZN
ZO ZO
ZP ZP
ZQ ZQ
ZR ZR
ZS ZS
ZT ZT
ZU ZU
ZV ZV
ZW ZW
ZX ZX
ZY ZY
ZZ ZZ

```

[illegible]

RESULT 377
ID N60695 standard: DNA: 2546 BP.
AC T62842;
DT 16-OCT-1991 (first entry)
DE Fragment pHS 53 homologous to human chromosome 11 pre-central region.
KW Homo sapiens.
RN Prenatal diagnosis; ss.
PD 07-JAN-1995
PF 02-AUG-1984; 778107.
PR 02-AUG-1984; SU-778107.
PT (MORSE), AUS PSYCHIKT HEALWH.
PT G0411, MW PSYCHIKT HEALWH.
DR WPI: 86-211201/32
PT Marking of eleventh human chromosome - involves using recombinant plasmid DNA plus 53 and genome DNA plus 53 fragment.
CC The probe is a 1.3 kb fragment of human chromosome 11 pre-central region of the human chromosome 11, useful in prenatal diagnosis, pathology etc.
SQ Sequence 2546 BP; 743 A; 493 C; 535 G; 785 T;
Query Match 1.3k; Score 14; DB 2; Length 2546;
Best Local Similarity 100.0%; Pred. No. 3.47e+01;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 594 tctgacatctctt 607
QY 941 TCAGAACTACTTT 954
RESULT 378
ID Q64889 standard: cDNA: 2546 BP.
AC Q64889;
DT 07-NOV-1994 (first entry)
DE DNA encoding human PPIase in insert pBS516.
KW PPIase; protein transferase; inhibition; framesylation; ss.
RN Homo sapiens.
FH Key Location/Qualifiers
FT cds 23..1345
FT /note= "alpha subunit"
FT cds 1348..2487
FT /note= "beta subunit"
FT W09410184-A.
PT 11-MAY-1994.
PD 29-OCT-1993; U10442.
PR 30-OCT-1992; DS-968782.
PA (MERF) MERCK & CO INC.
PI Diem RE, Gibbs JB, Kohl NE, Omer CA;
WPI: 94-167373/20.
DR P-PDSB; 854832.
PT Mammalian fattyacyl protein transferase proda. - used in assays for ops with inhibitory activity for the identification of PPIase.
PS Disclosure: Fig 5; 69pp: English.
CC The cDNA encoding the human alpha and beta subunits of PPIase was isolated from a human placental cDNA library in lambda gtl1 using a PPIase cDNA probe. The PPIase can be used to assess the inhibition of PPIase by various compounds and to identify anticancer agents.
CC See also Q64886-8.
SQ Sequence 2546 BP; 689 A; 607 C; 653 G; 597 T;
Query Match 1.3k; Score 15; DB 10; Length 2546;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 307 tctgacatctctt 321
Cp 450 TCTGATCCGAGCG 436
RESULT 379
ID T30738 standard: DNA: 2551 BP.
AC T30738;
DT 20-OCT-1996 (first entry)
DE Cryptdin-3, an anti-infective; antimicrobial; defense; inflammation; anti-inflammatory; inflammatory bowel disease; pancreatitis; cancer; tumour; ileitis; da.
KW Rattus sp.
RN Rattus sp.
FH Key Location/Qualifiers
FT exon 1..1482
FT /tag= a
FT /note= "exon 1 codes for the 5' untranslated region of cryptdin-3 prepro sequence"
FT intron 1483..2057
FT /tag= b
FT exon 2058..2551
FT /tag= c
FT /note= "exon 2 codes for cryptdin-3 and 3' untranslated region"
PT W09616075-A1.
PD 30-MAY-1996. U13328
PR 18-OCT-1994; DS-142268
PR 18-NOV-1994; DS-142268
PA (SHERI) SHRINER'S HOSPITAL FOR CRIPPLED CHILDREN.
PA (BREC) UNIV CALIFORNIA.
PI Ouellette AJ, Sealed ME;

DR P-PDSB; W14731.
PT New isolated osteogen receptor beta - used to develop proda. for treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular disease
PS Claim 6; Fig 1; 45pp: English.
CC related nuclear receptor (W1723), designated ER-beta. A set of degenerate primers based on conserved sequences of the DNA-binding domain and ligand binding domain of nuclear receptors were used to amplify rat prostate cDNA. The PCR prod. was used to identify rat ER-beta cDNA. The ER-beta cDNA was then used to identify human ER-beta cDNA (see also T62843)
CC This was then used to obtain human ER-beta cDNA (see also T62843) from a human testis cDNA library. Rat, human and mouse ER-beta (W14731-25) can be used to isolate molecules for use in the treatment of cardiovascular diseases, central nervous system diseases, osteoporosis, osteoarthritis, osteomyelitis, osteosarcoma, prostatic hyperplasia and to test environmental chemicals for osteogenic activity.
SQ Sequence 2568 BP; 634 A; 654 C; 663 G; 617 T;
Query Match 1.3k; Score 15; DB 29; Length 2568;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1840 actaacacacacag 1854
Cp 267 ACTAGACAAAGAG 253
RESULT 381
ID Q64237 standard: DNA: 2568 BP.
AC Q64237;
DT 28-JAN-1991 (first entry)
DE Human osteogen receptor beta (ER-beta) cDNA.
KW Enzyme; ss.
OS Acetobacter xylinum IPO 3288.
FH Key Location/Qualifiers
FT cds 815..1859
FT /tag= a
FT /product=cellulase
PT J02222688-A.
PD 05-SEP-1990. 045570
PR 27-FEB-1989; JP-045570
PR 27-FEB-1989; JP-045570
PA (SEER-) SHIN ENERGY GANYO.
PA (SEER-) KENYO-YO ALCOHOL KAIYATSU.
DR WPI: 90-31559/42.
PT Cellulase coding DNA sequence - can produce cellulase by bioengineering method
PT Claim 1; Fig 1; 7pp: Japanese.

DR P-PDSB; 898788;
PT New isolated cryptdin peptide(a) - which have antimicrobial activity in the detection and treatment of inflammatory pathologies
PS Claim 31; Page 65-66; 103pp: English.
CC The rat cryptdin-3 gene (T30738) codes for the precursor of cryptdin-3, an antimicrobial peptide.
CC T30738 encodes cryptdin-3, an antimicrobial peptide and an opportunistic pathogen.
CC Cryptdin-3, an antimicrobial peptide, is a member of the cryptdin family of proteins.
CC (see also T30736-37) were isolated from a Sprague-Dawley genomic library cloned in pEMBL3. The cryptdin-3 gene, or large amounts of cryptdin-3, can be used to produce pathologies of the intestine.
CC Pathologies of the intestine for use in treating inflammatory diseases.
SQ Sequence 2551 BP; 689 A; 543 C; 563 G; 756 T;
Query Match 1.3k; Score 15; DB 27; Length 2551;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1153 gactacatctctt 1167
Cp 660 GACATCAAGGAT 646
RESULT 380
ID T62842 standard: cDNA: 2568 BP.
AC T62842;
DT 08-JUN-1997 (first entry)
DE Osteogen receptor beta (ER-beta) cDNA.
KW nuclear receptor; prostate cancer; benign prostatic hyperplasia; ovary cancer; cardiovascular disease; osteoporosis; environment; pollutant; ss.
RN Rattus sp.
FH Key Location/Qualifiers
FT cds 424..1881
FT /tag= a
FT /note= "alternative start codon at 448..450"
PT W09709348-A2.
PD 13-MAR-1997.
PR 05-SEP-1996; E09333.
PR 05-SEP-1996; GB-018272.
PR 11-MAR-1996; GB-005550.
PR 11-MAR-1996; GB-005550.
PR 08-MAY-1996; GB-003576.
PA (KARO-) KARO BIO AB.
PI Emerck E, Gustafsson J, Kuiper GG;
WPI: 97-192842/1.


```

KW Rice, starch; transit peptide; pectin; cereal; amylopectin; seeds;
KW reverse transcriptase; plaques; ss.
OS Oryza sativa.
ST 5 utr
FT Location/Qualifiers
FT /tag= a
FT /tag= b
FT /tag= c
FT /tag= d
FT /tag= e
FT 1..2741
FT 1..197
FT 138..2477
FT 2478..2733
FT 3'utr
FT /tag= e
FT J05317057-A
FT P1
FT P2
FT P3
FT P4
FT P5
FT P6
FT P7
FT P8
FT P9
FT P10
FT P11
FT P12
FT P13
FT P14
FT P15
FT P16
FT P17
FT P18
FT P19
FT P20
FT P21
FT P22
FT P23
FT P24
FT P25
FT P26
FT P27
FT P28
FT P29
FT P30
FT P31
FT P32
FT P33
FT P34
FT P35
FT P36
FT P37
FT P38
FT P39
FT P40
FT P41
FT P42
FT P43
FT P44
FT P45
FT P46
FT P47
FT P48
FT P49
FT P50
FT P51
FT P52
FT P53
FT P54
FT P55
FT P56
FT P57
FT P58
FT P59
FT P60
FT P61
FT P62
FT P63
FT P64
FT P65
FT P66
FT P67
FT P68
FT P69
FT P70
FT P71
FT P72
FT P73
FT P74
FT P75
FT P76
FT P77
FT P78
FT P79
FT P80
FT P81
FT P82
FT P83
FT P84
FT P85
FT P86
FT P87
FT P88
FT P89
FT P90
FT P91
FT P92
FT P93
FT P94
FT P95
FT P96
FT P97
FT P98
FT P99
FT P100
FT P101
FT P102
FT P103
FT P104
FT P105
FT P106
FT P107
FT P108
FT P109
FT P110
FT P111
FT P112
FT P113
FT P114
FT P115
FT P116
FT P117
FT P118
FT P119
FT P120
FT P121
FT P122
FT P123
FT P124
FT P125
FT P126
FT P127
FT P128
FT P129
FT P130
FT P131
FT P132
FT P133
FT P134
FT P135
FT P136
FT P137
FT P138
FT P139
FT P140
FT P141
FT P142
FT P143
FT P144
FT P145
FT P146
FT P147
FT P148
FT P149
FT P150
FT P151
FT P152
FT P153
FT P154
FT P155
FT P156
FT P157
FT P158
FT P159
FT P160
FT P161
FT P162
FT P163
FT P164
FT P165
FT P166
FT P167
FT P168
FT P169
FT P170
FT P171
FT P172
FT P173
FT P174
FT P175
FT P176
FT P177
FT P178
FT P179
FT P180
FT P181
FT P182
FT P183
FT P184
FT P185
FT P186
FT P187
FT P188
FT P189
FT P190
FT P191
FT P192
FT P193
FT P194
FT P195
FT P196
FT P197
FT P198
FT P199
FT P200
FT P201
FT P202
FT P203
FT P204
FT P205
FT P206
FT P207
FT P208
FT P209
FT P210
FT P211
FT P212
FT P213
FT P214
FT P215
FT P216
FT P217
FT P218
FT P219
FT P220
FT P221
FT P222
FT P223
FT P224
FT P225
FT P226
FT P227
FT P228
FT P229
FT P230
FT P231
FT P232
FT P233
FT P234
FT P235
FT P236
FT P237
FT P238
FT P239
FT P240
FT P241
FT P242
FT P243
FT P244
FT P245
FT P246
FT P247
FT P248
FT P249
FT P250
FT P251
FT P252
FT P253
FT P254
FT P255
FT P256
FT P257
FT P258
FT P259
FT P260
FT P261
FT P262
FT P263
FT P264
FT P265
FT P266
FT P267
FT P268
FT P269
FT P270
FT P271
FT P272
FT P273
FT P274
FT P275
FT P276
FT P277
FT P278
FT P279
FT P280
FT P281
FT P282
FT P283
FT P284
FT P285
FT P286
FT P287
FT P288
FT P289
FT P290
FT P291
FT P292
FT P293
FT P294
FT P295
FT P296
FT P297
FT P298
FT P299
FT P300
FT P301
FT P302
FT P303
FT P304
FT P305
FT P306
FT P307
FT P308
FT P309
FT P310
FT P311
FT P312
FT P313
FT P314
FT P315
FT P316
FT P317
FT P318
FT P319
FT P320
FT P321
FT P322
FT P323
FT P324
FT P325
FT P326
FT P327
FT P328
FT P329
FT P330
FT P331
FT P332
FT P333
FT P334
FT P335
FT P336
FT P337
FT P338
FT P339
FT P340
FT P341
FT P342
FT P343
FT P344
FT P345
FT P346
FT P347
FT P348
FT P349
FT P350
FT P351
FT P352
FT P353
FT P354
FT P355
FT P356
FT P357
FT P358
FT P359
FT P360
FT P361
FT P362
FT P363
FT P364
FT P365
FT P366
FT P367
FT P368
FT P369
FT P370
FT P371
FT P372
FT P373
FT P374
FT P375
FT P376
FT P377
FT P378
FT P379
FT P380
FT P381
FT P382
FT P383
FT P384
FT P385
FT P386
FT P387
FT P388
FT P389
FT P390
FT P391
FT P392
FT P393
FT P394
FT P395
FT P396
FT P397
FT P398
FT P399
FT P400
FT P401
FT P402
FT P403
FT P404
FT P405
FT P406
FT P407
FT P408
FT P409
FT P410
FT P411
FT P412
FT P413
FT P414
FT P415
FT P416
FT P417
FT P418
FT P419
FT P420
FT P421
FT P422
FT P423
FT P424
FT P425
FT P426
FT P427
FT P428
FT P429
FT P430
FT P431
FT P432
FT P433
FT P434
FT P435
FT P436
FT P437
FT P438
FT P439
FT P440
FT P441
FT P442
FT P443
FT P444
FT P445
FT P446
FT P447
FT P448
FT P449
FT P450
FT P451
FT P452
FT P453
FT P454
FT P455
FT P456
FT P457
FT P458
FT P459
FT P460
FT P461
FT P462
FT P463
FT P464
FT P465
FT P466
FT P467
FT P468
FT P469
FT P470
FT P471
FT P472
FT P473
FT P474
FT P475
FT P476
FT P477
FT P478
FT P479
FT P480
FT P481
FT P482
FT P483
FT P484
FT P485
FT P486
FT P487
FT P488
FT P489
FT P490
FT P491
FT P492
FT P493
FT P494
FT P495
FT P496
FT P497
FT P498
FT P499
FT P500
FT P501
FT P502
FT P503
FT P504
FT P505
FT P506
FT P507
FT P508
FT P509
FT P510
FT P511
FT P512
FT P513
FT P514
FT P515
FT P516
FT P517
FT P518
FT P519
FT P520
FT P521
FT P522
FT P523
FT P524
FT P525
FT P526
FT P527
FT P528
FT P529
FT P530
FT P531
FT P532
FT P533
FT P534
FT P535
FT P536
FT P537
FT P538
FT P539
FT P540
FT P541
FT P542
FT P543
FT P544
FT P545
FT P546
FT P547
FT P548
FT P549
FT P550
FT P551
FT P552
FT P553
FT P554
FT P555
FT P556
FT P557
FT P558
FT P559
FT P560
FT P561
FT P562
FT P563
FT P564
FT P565
FT P566
FT P567
FT P568
FT P569
FT P570
FT P571
FT P572
FT P573
FT P574
FT P575
FT P576
FT P577
FT P578
FT P579
FT P580
FT P581
FT P582
FT P583
FT P584
FT P585
FT P586
FT P587
FT P588
FT P589
FT P590
FT P591
FT P592
FT P593
FT P594
FT P595
FT P596
FT P597
FT P598
FT P599
FT P600
FT P601
FT P602
FT P603
FT P604
FT P605
FT P606
FT P607
FT P608
FT P609
FT P610
FT P611
FT P612
FT P613
FT P614
FT P615
FT P616
FT P617
FT P618
FT P619
FT P620
FT P621
FT P622
FT P623
FT P624
FT P625
FT P626
FT P627
FT P628
FT P629
FT P630
FT P631
FT P632
FT P633
FT P634
FT P635
FT P636
FT P637
FT P638
FT P639
FT P640
FT P641
FT P642
FT P643
FT P644
FT P645
FT P646
FT P647
FT P648
FT P649
FT P650
FT P651
FT P652
FT P653
FT P654
FT P655
FT P656
FT P657
FT P658
FT P659
FT P660
FT P661
FT P662
FT P663
FT P664
FT P665
FT P666
FT P667
FT P668
FT P669
FT P670
FT P671
FT P672
FT P673
FT P674
FT P675
FT P676
```

```

Tue Nov 17 08:55:26 1998          US-08-087-977-9.mmg

misc_difference 233...239
      /tag= c
      /note= "bases n at positions 233-240 are unclear
              in the specification"

C3X1080606-N.
PD 16-APR-1995.
PF 15-OCT-1993; 108606.
PP 15-OCT-1993; CA-108606.
PR (JIKH) JINSEI NAKA SAKUTE & REICH MEDICALE.
PA (JIKH) JINSEI NAKA SAKUTE J. Mosser J., Serde C.
PI WP: 95-215721/29.
PQ P-PSDB: R76110.
DR PR New nucleic acid responsible for adreno-leuco-dystrophy - related
    PT proteins and antibodies, useful for diagnosis and treatment
    PP probes corresp. to fragments in the red pigment gene of an ANK
    PQ protein, expressed in Escherichia coli, deposited as a deposit
    CC following hybridizations and cloned PCR, a HeLa cell cDNA library
    CC was screened, and a cDNA fragment contg. the complete coding
    CC sequence of the gene was obtained. The gene encodes a transmembrane
    CC hemopoietic cells with the sequence in vivo will allow therapy
    CC of ALD or ANK.
    DQ Sequence 2746 BP: 494 A: 964 C: 815 G: 466 T:
        Query Match 130. Score 14: DB 15: Length 2746;
        Best Local Similarity 10.38: Pred No 3 47+02:
        Matches 14: Conservative O: Mismatches G: Indels O: Gaps

Db 640 tctgttcggcagg 653
OY 1012 TCTGTTCGGCAGG 1025

RESULT_394
ID T193315 standard; cDNA, 2750 BP.
AC T193315;
CD T193315;
DE 26-FEB-1997 (first entry)
DT Human adrenoleucodystrophy (ALD) wild-type gene.
KW Adrenoleucodystrophy; gene therapy; retroviral vector M49;
KW Adrenoleucodystrophy; membrane protein; long chain fatty acid oxidation
KW Adrenoleucodystrophy; membrane protein; long chain fatty acid oxidation
KW Homo sapiens.
FH Key
FH cds
FT TT Location/Qualifiers
FT TT 387..2524
FT TT /product= membrane_protein
FT TT /tag= b
FT TT 1221..1829
FT TT /label= Ex13
FT TT /note= "probe Ex13 corresponds to this region."
FT TT W0961273-A2.

```

PD 18-JUL-1996. PRO058.
 PR 13-JAN-1995. PR-00376.
 PI (INRM) INST NAT SANTE & RECH MEDICALE.
 PI Aubourg P, Mandel JL, Mosser J, Sarda CO;
 DR WPI: 88-30136/43.
 PT Recombinant viral vector contg. DNA for correcting
 PT adrenoleuco-dystrophy - and immunologically tolerable cells contg.
 PT this vector, useful in gene therapy
 CC A family of related DNA insertion sequences (ISM) in the genome of
 CC disease isolates of mycobacteria has been isolated. The first example of
 CC ISM (ISM-1) was identified and characterized in the clone pMB22 (NCIB
 CC 10240) derived from Mycobacterium fortuitum. The ISM-1 was identified
 CC from Crohn's disease. The ISM-1 was also identified in Mycobacterium
 CC paratuberculosis cultured from cattle and other animals suffering from
 CC Johne's disease. ISM0976 gives the sequence of the portion of pMB22
 CC which spans the ISM-1 insertion element. The DNA can be used as DNA or
 CC probe to identify ISM-1 insertion elements and to study the genetic
 CC mycobacteria in environmental and disease samples and in vitro
 CC closely related mycobacterial strains. The DNA probes allow discrimination between
 CC can be used in diagnostics or for producing antibodies or as vaccines
 CC Sequence 2762 BP; 432 A; 897 G; 918 G; 513 T; 2 Others;
 SQ Sequence 2762 BP; 432 A; 897 G; 918 G; 513 T; 2 Others;

Query Match 1.38; Score 14; DB 25; Length 2762;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 643 tctgtgctggag 656
 QY 1012 TCTGTGCTGGAG 1025

RESULT 395
 ID N80976 standard; DNA: 2762 BP.
 DC 12-SEP-1990 (first entry)
 DE contg. 2 contiguous lengths of DNA (contigs C1 and C2) on pMB22
 DE contg. Mycobacterial insertion sequences (ISM)
 KW Mycobacterial insertion sequences; Mycobacterium; pMB22; NCIB 12461;
 KW Mycobacterial strain discrimination; vaccine;
 KW Mycobacterial disease prevention; as
 OS Mycobacterium.
 FH Key Location/Qualifiers
 FT misc_feature 1..1280
 FT cds 1280..1724
 FT /note="contig C1"
 FT misc_feature 1261..2762
 FT /tag="b"
 FT /note="contig C2"
 EP-288306-A.
 PD 26-OCT-1988.

PF 22-APR-1988. 303657.
 PR 24-APR-1987. GB-009803.
 PI (MCPA) McFadden J.
 PI C. J. McFadden, Jameson-Taylor J;
 DR WPI: 88-30136/43.
 PT DNA insertion sequences of mycobacterial origin -
 PT producing proteins for vaccines, immuno therapy or diagnostics
 CC A family of related DNA insertion sequences (ISM) in the genome of
 CC disease isolates of mycobacteria has been isolated. The first example of
 CC ISM (ISM-1) was identified and characterized in the clone pMB22 (NCIB
 CC 10240) derived from Mycobacterium fortuitum. The ISM-1 was identified
 CC from Crohn's disease. The ISM-1 was also identified in Mycobacterium
 CC paratuberculosis cultured from cattle and other animals suffering from
 CC Johne's disease. ISM0976 gives the sequence of the portion of pMB22
 CC which spans the ISM-1 insertion element. The DNA can be used as DNA or
 CC probe to identify ISM-1 insertion elements and to study the genetic
 CC mycobacteria in environmental and disease samples and in vitro
 CC closely related mycobacterial strains. The DNA probes allow discrimination between
 CC can be used in diagnostics or for producing antibodies or as vaccines
 CC Sequence 2762 BP; 432 A; 897 G; 918 G; 513 T; 2 Others;
 SQ Sequence 2762 BP; 432 A; 897 G; 918 G; 513 T; 2 Others;

Query Match 1.38; Score 14; DB 1; Length 2762;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2412 tgaatgtgggctt 2425
 QY 611 TGAATGTGGGCTT 624

RESULT 396
 ID N80976 standard; DNA: 2763 BP.
 DC 12-SEP-1990 (first entry)
 DE 27-MAR-1998 (first entry)
 DE Human disintegrin metalloprotease (aggrinase) gene.
 KW Disintegrin; metalloprotease; arthritis; osteoarthritis; diagnosis;
 KW aggrinase; arthropathy; osteoporosis; ankylosing spondylitis;
 KW Crohn's disease; human; as.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 17..2416
 FT /tag="a"
 PD 04-SEP-1987.
 PF 28-FEB-1997; U03217.
 PR 01-MAR-1996; US-013679.
 PR (PROC) PROCTER & GAMBLE CO.

PA (UYCA) UNIV CASE WESTERN RESERVE.
 PI Haqvi T, Tindal MH;
 PR 27-SEP-1997. 44826/41.
 PI New human disintegrin metalloprotease - is differentially expressed
 PT in arthritis, used to screen for drugs and to diagnose
 PT metalloprotease-mediated diseases, e.g. osteoarthritis
 CC This gene encodes a human disintegrin metalloprotease that is
 CC up-regulated in human chondrocytes during osteoarthritic conditions
 CC and which is involved in tissue remodeling and breakdown. It was
 CC obtained by screening human cDNA libraries with a partial cDNA
 CC stimulated human articular (femoral head) chondrocytes using
 CC differential display analysis. Disintegrin nucleic acids can be
 CC used in the construction of recombinant expression vectors for
 CC a marker to monitor the development of arthritis. The gene provides
 CC asymptomatic stages, and the progression of osteoarthritis.
 CC Inhibitors of the disintegrin, including antisense oligonucleotides,
 CC can potentially be used in the treatment of metalloprotease-mediated
 CC spondylitis, rheumatoid or inflammatory arthritis, gout and Lyme
 CC disease.
 SQ Sequence 2763 BP; 849 A; 537 C; 610 G; 767 T;

Query Match 1.38; Score 14; DB 36; Length 2763;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2487 agtcatcatcca 2500
 QY 498 AGTCATCATCCA 511

RESULT 397
 ID Q13239 standard; cDNA: 2781 BP.
 AC Q13239.
 DT 29-OCT-1991 (first entry)
 DE Heat shock factor; as.
 KW Drosophila.
 FH Key Location/Qualifiers
 FT cds 239..2304
 FT /tag="a"
 FT poly_a_signal 2729..2727
 FT /tag="b"
 FT poly_a_site 1757..1781
 FT /tag="c"
 PD 087617901-A.

PD 16-JUL-1991. 617901.
 PR 26-NOV-1990. US-617901.
 PI (USSH) NAT INST OF HEALTH
 PI C. J. McFadden, Jameson-Taylor J, Rabindran S;
 DR WPI: 91-25234/31.
 DR P-PDSB: R13502.
 PT DNA encoding Drosophila and human heat shock factor proteins -
 PT used for developing prods. for studying stress and disease states -
 CC The sequence encodes Drosophila heat shock factor protein and was
 CC obtained by screening a Drosophila genomic library with oligo-
 CC nucleotide probes (Q13237, Q13238) based on the HSF amino acid
 CC sequence. The protein is a member of the HSF family of proteins in
 CC other organisms and also for the detection of stress or a diseased
 CC state in living systems. The gene can be used to increase
 CC expression of other gene prods. by cotransfecting the HSF gene
 CC into cells with other gene products. Heat shock elements it can be
 CC linked to tissue-specific promoters and used to study gene expression
 CC or introduced into transgenic mice as a tool for eliciting increased
 CC or chronic stress response conditions as a model for how tissues
 CC respond to chronic stress conditions such as those caused by viral
 CC infection, chemical or mechanical stress. See also Q13240 and
 CC Q13241.
 SQ Sequence 2781 BP; 831 A; 631 C; 690 G; 629 T;

Query Match 1.38; Score 14; DB 2; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 990 gaccagtggagcc 1003
 QY 1050 GACCAGTGGAGCC 1063

RESULT 398
 ID Q13239 standard; DNA: 2800 BP.
 AC Q13239.
 DT 15-JUL-1994 (first entry)
 DE Meningitis IM2394 transferrin receptor Tbp1 subunit DNA sequence.
 KW Meningitis IM2394 transferrin receptor Tbp1 subunit; meningitis vaccine;
 KW Meningitis IM2394 transferrin receptor Tbp1 subunit; meningitis vaccine;
 OS Neisseria meningitidis (strain IM2394).
 FH Key Location/Qualifiers
 FT cds 40..2766
 FT /tag="a"
 FT signal_peptide 40..111
 FT /tag="b"
 FT mat_peptide 112..2763
 FT /tag="c"

Tue Nov 17 08:55:26 1998

```

Query Match      1.3%  Score 15; DB 2; Length 2876;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

b 734 GGGTAAAGGAACCA 768
  |||||
P 650 GGGTAAAGGAACCA 636

```

RESULT	407	D	GB4613 standard; cDNA, 2884 BP.	
C	GB4613;			
C	04-SEP-1995		(first entry)	
C	04-SEP-1995		1951-1952	
C	04-SEP-1995		1951-1952	
M	AT2 receptor; angiotensin II type 2 receptor; transgenic animal.			
M	gene therapy; vector; antibody; diagnosis; ds.			
S	status sp.			
S	polya_signal		Location/Qualifiers	
T	2234..2239		2234..2239	
T	polya_signal		2831..2836	
T	polya_signal		/'tag= a	
T	polya_signal		/'tag= b	
D	N09504073-A.			
D	09-FEB-1995.			
D	30-JUL-1994		U08055.	
R	30-JUL-1994		US-100486.	
R	30-JUL-1994		US-100486.	
R	30-JUL-1994		US-100486.	
R	30-JUL-1994		US-100486.	
I	(SPRD) UNIV ILELAND STANFORD JUNIOR.			
I	DEAU V, MUKOYAMA M.			
R	WP1: 95-082181/11.			
R	WP1: 95-082181/11.			
T	New mammalian angiotensin II type 2 receptor - and related			
T	nucleic acid, vectors, antibodies and transformed cells useful			
T	for identifying cpds. for treating and diagnosing AT2 related			
T	disorders.			
S	Abstracts: Page 32-34; 48pp; English.			
S	A fetal rat cDNA library in pCDNA3 was used to transform COS-7			
S	cells, and recombinants were tested for binding to 125I-labeled			
C	AT2 analogs. An active pool was screened further to isolate clones			
C	expressing the AT2 receptor gene for gene therapy. (GB4613). The			
C	AT2 receptor gene may be useful for gene therapy.			
C	Sequence 2884 BP: 819 A: 557 C: 542 G: 966 T:			

```

Query Match      1.38;      Score 14; DB 14; Length 2894;
Best Local Similarity 100.0%;      Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

b 2685 ttttgggttggaaga 2698
|||||
p 537 tttttgttggaaga 524

```

Tue Nov 17 08:55:26 1998

N	[15:17] (q21;q11-22)]: breakpoint; polymerase chain reaction; as.
S	Synthetic.
H	T
T	Key
C	Location/Qualifiers
C	103-1488
C	/tag= a
D	N09216660-A
D	Q1-OCT-1992.
D	UO3320.
R	23-MAR-1992: US-67368.
R	22-MAR-1991: US-67368.
R	23-MAR-1992: US-67368.
I	(SLOK) SZYAN KETTERING INST CANCER.
I	Dmitrovsky E, Evans RM, Frankel S, Kasizuka A, Miller WH;
I	Watrell BP, 1992,042.
I	P-SPB3 - 22755/42.
R	P-SPB3 - 22755/42.
T	Marker for acute promyelocytic leukaemia and other neoplasia -
T	comprising nucleic acid and encoded abnormal retinoic acid
T	receptor-alpha receptor 84pp; English
T	The sequence given represents the nucleic acid sequence of the
C	retinoic acid receptor (RAR)-alpha gene. This gene is disrupted in
C	a translocation of a portion of the long arm of chromosome 17 onto
C	the long arm of chromosome 15 t(15;17)(q21;q11-22); this causes a
C	promyelocytic leukemia (APL). The breakpoint region has been cloned
C	and it has been shown that DNA rearrangements are clustered in the
C	region of the first intron of RAR-alpha. This sequence was isolated
C	from a patient with APL. This sequence may be used to identify the
C	of this sequence. 582 bp used to amplify the translocated region.
C	Sequence 2928 BP: 592 A: 1012 C: 789 G: 545:
Q	

```

Query Match      1.3%   Score 15; DB 5; Length 2928;
Best Local Similarity 100.0%;   Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

ID	RESULT 408
ID	T66421 standard; cDNA to mRNA; 2922 BP.
AC	T66421
DE	766421; 0897 (first entry)
DE	Homo sapiens; P100 protein sequence.
DE	Homo P100 proteasome; P100 protein; protease; yeast; oncogene;
KW	human; 26S proteasome; P100 protein; protease; yeast; oncogene;
KW	cancer cell; liver; kidney; diagnosis; malignant tumour; da.
OS	Homo sapiens.
OS	Locations/Qualifiers
CD	34..2760
FT	/*tag= a
FT	/product= p100 protein
FT	J09071002.A.
FT	28-MAR-1997.
PP	11-SEP-1995; 232383.
PR	11-SEP-1995; JP-232383.
PA	(KIM) SANG-HO KIMOTO KK.
PA	(KIM) SANG-HO KIMOTO KK.
WP	WPI: 97-239284/22.
P-PDB:	P-PDB: W15473.
DR	Human 26S proteasome component P100 protein - acts to decompose
DR	protein with binding protein, for use in diagnosis and treatment of
CC	cancer.
CC	Claim 3, Page 12-13; 15pp; Japanese.
CC	This sequence encodes the human 26S proteasome component P100
CC	proteasome. The proteasome is a multifunctional protease ubiquitous
CC	in eukaryotic organisms. Proteasomes are composed of several
CC	decomposable oncogene products and a significant expression of a proteasome
CC	gene is known to occur in cancer cells in liver, kidney. Elucidation of
CC	the functions of the proteasome will contribute to the diagnosis and
CC	treatment of various malignant tumour in which
CC	proteasomes are thought to be involved.
CC	Sequence 2922 BP; 695 A; 698 C; 853 G; 876 T.
SQ	

```

Query Match      1.18; Score 14; DB 29; Length 2972;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 880 accctctgcaagga 893
Cp 94 ACCCTCTGCAAGGA 81

```

RESULT	409
ID	Q29338 standard; cDNA; 2928 bp.
AC	Q29338;
DT	09-MAR-1993 (first entry)
DD	RAR-alpha gene.
D2	Retinoic acid receptor; RAR-alpha; myl; acute promyelocytic leukemia;
KW	APL; translocation; chromosome 17; chromosome 15; PC9; primer:

Tue Nov 17 08:55:26 1998

cds	279..2255	
FF	/clog= a	
PD	W09A10187-A.	
PD	11-MAY-1994.	
PD	26-OCT-1993.	U10307.
PD	26-OCT-1993.	U10307.
PD	30-OCT-1992.	US-565914.
PD	(CNSJ) CONSEJO INVESTIGACION.	
PD	(HSCR) HSC RES & DEV LP.	
PD	(SLK) SLKON GENETICS INC.	
PD	WPI: 94-167376/720.	
PD	P-PDB: 854828.	
PD	Nucleic acid encoding endoglin or fragments - used for modifying	
PD	transforming growth factor-beta for	
PD	treating pathological conditions.	
PD	Claim 1; Fig 13; 60pp.	
PD	The cDNA encoding endoglin was isolated by PCR using a cDNA library	
PD	prepared from PMA-treated myelomonocytic human cell line H660	
PD	and the sequence of the cDNA was determined. The cDNA encodes the	
PD	regulatory function of TGF-beta. It can be used to treat	
PD	conditions caused by TGF-beta regulated cell growth stimulation. TGF-	
PD	beta regulated inhibition of cell growth or TGF-beta regulated	
PD	extracellular matrix stimulation.	
PD	See also Q64894-1.	
PD	Sequence 2937 BP; 588 A; 1034 C; 801 G; 514 T;	
PD	Query Match 1.33; Score 14; DB 100; Length 2937;	
PD	Matches 10; Similarity 100.0%; 0; Mismatches 0; Gaps	
PD	14; Conservative	

Db 1110 tctcttcagatctt 1123
|||||
Cp 971 tctcttcagatctt 988

RESULT 411
ID Q61476 standard: cDNA to RNA. 2940 bp.

ISSUES	411
STANDARD	G61476 standard; CDNA to mRNA, 2940 BP.
ID	Q81476 AC
DE	Q1-SEP-1995 (first entry)
EC	RAR-alpha clone lambda-4ir.
EW	RAR-alpha clone lambda-4ir.
KW	differentiation, da
OR	Homo sapiens.
Key	Location/Qualifiers
PF	chr.:14q1
CD	/seg= d
FT	
MO	M09504143-A.
PD	08-FEB-1995.
ED	08-FEB-1995.
EF	28-JUL-1994.
EE	21-MAY-1994.
EN	08-APR-1994.

QY 833 CTTGCCAGAGGAA 846

RESULT 426

ID T79696 standard: cDNA; 3214 BP.
AC 17-DEC-1997; 169479
DE Human origin of replication complex protein 1 gene.
DE KW Origin of replication complex; ORC; yeast; human; chromatography;
DE PE Peptide sequencing; primer; amplification; PCR; genome; cell growth;
DE KW Polymerase chain reaction; open reading frame; cancer; infection;
DE OS Homo sapiens. hyperplasia; da.
DE FH Key Location/Qualifiers
DE FT cde 220..2805
DE FT /tag= a
DE FT /product= ORC1 protein
DE PN US5614618-A.
DE PD 25-MAR-1997; 169479
DE PF 16-DEC-1997; 169479
DE PR 07-JUN-1995; US-484106
DE PA (COLD-) COLD SPRING HARBOR LAB.
DE PI (RESC) UNIV CALIFORNIA.
DE PI Cell Spn; Foss M; Gavin K; Herskowitz I; Hidaba M;
DE PI Stillman RW; Laurenceon F; Li J; McNally FJ; Rine J;
DE WI: 97-201534/48.
DE DR P-PSDB; W22232.
DE PT Nucleic acids encoding origin of replication complex proteins - used
DE PT associated with undesirable cell growth
DE PT This is the nucleotide sequence encoding the human origin of
DE PT replication complex protein 1 gene (ORC1) from the human genome.
DE CC between the 8. cerevisiae (W22224), K. lactis (W22230) and S. pombe
DE CC (W22232) proteins. The amplified fragment was then used to screen
DE CC a phage lambda cDNA library constructed from MZRD21, an embryonic
DE CC (W22224-35) can be used to screen chemical libraries to identify lead
DE CC compounds useful in treatment and diagnosis of undesired cell growth.
DE CC e.g. cancer, infections, inflammation and hypersensitivity.
DE CC Sequence 3214 BP; 897 A; 808 C; 792 G; 717 T;
Query Match 1.38; Score 14; DB 31; Length 3214;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2190 cttgccagagagaa 2203
|||||

QY 833 CTTGCCAGAGGAA 846

RESULT 425

ID T79696 standard: DNA; 3233 BP.
AC 17-DEC-1997; 169479
DE Human origin of replication complex protein 1 gene.
DE KW Origin of replication complex; ORC; yeast; human; chromatography;
DE PE Peptide sequencing; primer; amplification; PCR; genome; cell growth;
DE KW Polymerase chain reaction; open reading frame; cancer; infection;
DE OS Homo sapiens. hyperplasia; da.
DE FH Key Location/Qualifiers
DE FT cde 220..2805
DE FT /tag= a
DE FT /product= ORC1 protein
DE PN US5614618-A.
DE PD 25-MAR-1997; 169479
DE PF 16-DEC-1997; 169479
DE PR 07-JUN-1995; US-484106
DE PA (COLD-) COLD SPRING HARBOR LAB.
DE PI (RESC) UNIV CALIFORNIA.
DE PI Cell Spn; Foss M; Gavin K; Herskowitz I; Hidaba M;
DE PI Stillman RW; Laurenceon F; Li J; McNally FJ; Rine J;
DE WI: 97-201534/48.
DE DR P-PSDB; W22232.
DE PT Nucleic acids encoding origin of replication complex proteins - used
DE PT associated with undesirable cell growth
DE PT This is the nucleotide sequence encoding the human origin of
DE PT replication complex protein 1 gene (ORC1) from the human genome.
DE CC between the 8. cerevisiae (W22224), K. lactis (W22230) and S. pombe
DE CC (W22232) proteins. The amplified fragment was then used to screen
DE CC a phage lambda cDNA library constructed from MZRD21, an embryonic
DE CC (W22224-35) can be used to screen chemical libraries to identify lead
DE CC compounds useful in treatment and diagnosis of undesired cell growth.
DE CC e.g. cancer, infections, inflammation and hypersensitivity.
DE CC Sequence 3214 BP; 897 A; 808 C; 792 G; 717 T;
Query Match 1.38; Score 14; DB 31; Length 3214;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2190 cttgccagagagaa 2203
|||||

CC diagnosing susceptibility or predisposition to cancer in a patient. The
CC nucleic acid molecules are used to design probes or primers for PCR to
CC determine or detect the presence of mutations in a sample of nucleic
CC acid. The probes or primers are used to amplify the nucleic acid under
CC conditions which modulate the expression of nucleic acid under
CC control of the promoter. Antibodies are used to determine the presence,
CC amount or location in a cell of a BRCA2 polypeptide or its mutant forms.
CC The polypeptides are used to screen for binding partners, these are
CC polypeptides which can be used as cancer therapeutics.
CC Sequence 3233 BP; 1039 A; 592 C; 556 G; 1046 T;
Query Match 1.38; Score 15; DB 15; Length 3233;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3058 taagttgatgcat 3072
CP 361 TAAAGTTCATGTCAT 347
|||||

QY 833 CTTGCCAGAGGAA 846

RESULT 427

ID T79696 standard: DNA; 3103 BP.
AC 17-DEC-1997; 169479
DE Human origin of replication complex protein 1 gene.
DE KW Origin of replication complex; ORC; yeast; human; chromatography;
DE PE Peptide sequencing; primer; amplification; PCR; genome; cell growth;
DE KW Polymerase chain reaction; open reading frame; cancer; infection;
DE OS Homo sapiens. hyperplasia; da.
DE FH Key Location/Qualifiers
DE FT cde 220..2805
DE FT /tag= a
DE FT /product= ORC1 protein
DE PN US5614618-A.
DE PD 25-MAR-1997; 169479
DE PF 16-DEC-1997; 169479
DE PR 07-JUN-1995; US-484106
DE PA (COLD-) COLD SPRING HARBOR LAB.
DE PI (RESC) UNIV CALIFORNIA.
DE PI Cell Spn; Foss M; Gavin K; Herskowitz I; Hidaba M;
DE PI Stillman RW; Laurenceon F; Li J; McNally FJ; Rine J;
DE WI: 97-201534/48.
DE DR P-PSDB; W22232.
DE PT Nucleic acids encoding origin of replication complex proteins - used
DE PT associated with undesirable cell growth
DE PT This is the nucleotide sequence encoding the human origin of
DE PT replication complex protein 1 gene (ORC1) from the human genome.
DE CC between the 8. cerevisiae (W22224), K. lactis (W22230) and S. pombe
DE CC (W22232) proteins. The amplified fragment was then used to screen
DE CC a phage lambda cDNA library constructed from MZRD21, an embryonic
DE CC (W22224-35) can be used to screen chemical libraries to identify lead
DE CC compounds useful in treatment and diagnosis of undesired cell growth.
DE CC e.g. cancer, infections, inflammation and hypersensitivity.
DE CC Sequence 3233 BP; 1039 A; 592 C; 556 G; 1046 T;
Query Match 1.38; Score 14; DB 24; Length 3231;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2679 tcttcgaagagtag 2692
CP 961 ATTCCAGAGGAGG 948
|||||

PR 14-MAY-1993; US-062443.
PA (UVA), UNIV. JEFFERSON THOMAS.
PI Cannani E, Croce C.
DR NPI: 95-066818/01.
DR P-PDB: R66452.
CC A few acute lymphocytic leukemia gene products - used for the
CC detection of acute lymphocytic leukemia.
CC Lymphoblastic or nonlymphoblastic leukemia
CC
CC Claim 8, Page 110-111: 207pp; English.
CC Oligonucleotide probes of at least 15 nucleotides derived from
CC the sequence of the human integrin beta subunit gene are claimed.
CC The sequence within the AP-9 encodes a nuclear targeting sequence
CC the protein encoded by AP-9 encodes a nuclear targeting sequence
CC AKKOK. The AP-9 protein is serine-rich (20%) including an
CC uninterrupted stretch of 42 serine residues and has a higher than
CC average pI.
SQ Sequence 3176 BP; 1083 A; 713 C; 698 G; 882 T;
Query Match
Best Local Similarity 100.0%; Pred. No. 6.37e+01; Length 3376;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 312 ggaagcagatcat 326
Cp 93 GAGACGAGTACAT 69

RESULT 432
ID T36481 standard; cDNA: 3415 BP.
DT 08-OCT-1996 (first entry)
DE Human integrin beta subunit protein, beta-5, cDNA.
KW Human integrin beta subunit; beta-5; carcinoma; lymphoid cell;
KW Immunodetection; detection; mRNA; assay; ss.
PR Homo sapiens. Location/Qualifiers
cfa 337..2736
FI signal_peptide 337..405
FI mat_peptide 406..2733
FI /tag= a
FI /tag= c
FI US527679-A.
PR 19-JUN-1996. 694314.
PR 01-MAY-1991. US-694314.
PR 27-APR-1993: US-054077.
PA (DAND) DNA FARMER CANCER INST INC.
PI Hemley ME, Rajaswamy H.
DR P-PDB: R62194.

Db 1439 tgaagatttgag 1452
Cp 956 TGAAGATTGAG 969

RESULT 434
ID T80099 standard; cDNA: 3423 BP.
DT 28-NOV-1997 (first entry)
DE Human NF-1 gene.
KW Alzheimer's disease; NF-1; human; mouse; NF361 gene; antibody; ss.
PR Homo sapiens. Location/Qualifiers
cfa 923..2380
FI /tag= a
FI J09215495-A.
PR 30-JUN-1996. 181514.
PR 07-DEC-1995: JP-345659.
PA (SUNG) SUMITOMO SEITAKU KK.
DR NPI: 97-455073/43.
PR A human NF1 gene expression gene - used in the diagnosis of
PR Alzheimer's disease
PS Claim 2, Page 17-18: 23pp; Japanese.
CC This sequence represents the human NF1 gene. This sequence contains the
CC 5' and 3' untranslated regions (UTRs) and the coding region. The
CC T80100 all represent cDNA sequences of the invention. The sequences are
CC all NF-1 genes, and contain the open reading frames shown in T80103.
CC T80102 and T80101 respectively. These genes, their products and
CC antibodies specific for the proteins are useful for the diagnosis of
CC Alzheimer's disease.
SQ Sequence 3423 BP; 811 A; 808 C; 867 G; 937 T;
Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 3423;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2568 gtgaagagatcat 2581
Cp 982 GTGAGAGAGATCA 995

RESULT 435
ID Q57072 standard; DNA: 3425 BP.
DT 22-JUL-1994 (first entry)
DE AGE-modified DNA INS-8.
KW Advanced glycosylation end products; AGE plasmids; transposon; ss.
OS Mus musculus.

PR Immunodetection assay for beta-5 protein - useful
PR for the detection of carcinoma(s) and to distinguish different cell
PR types.
PS Claim 11: Columns 13-20: 21pp; English.
CC The present sequence encodes the human integrin beta subunit
CC protein, beta-5, which is found in carcinomas but not in lymphoid
CC cells. This sequence is useful for the detection of beta-5, comprises a
CC conserved immunodetection site, a cytosolic domain, and a transmembrane
CC epitope of beta-5, and then with a labelled Ab which binds another
CC epitope of beta-5 and detecting any bound label. An assay for
CC beta-5 mRNA, comprises contacting a sample with a probe capable of
CC binding to the beta-5 cDNA and determining if binding has
CC occurred. These assays are useful for detecting carcinomas, and for
CC distinguishing between different cell types.
SQ Sequence 3415 BP; 752 A; 943 C; 1009 G; 711 T;
Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 3415;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1869 cttctctctctctct 1882
Cp 1013 CTTCTCTCTCTCTCT 1026

RESULT 433
ID T75437 standard; DNA: 3420 BP.
DT 07-OCT-1997 (first entry)
DE Mortierella alpina cytochrome b5 genomic DNA.
KW Mortierella alpina; cytochrome b5; preparation; production; human;
KW fatty acids; ss.
PR Mortierella alpina.
PR J09121873-A.
PR 11-MAY-1997. 203735.
PR 01-AUG-1995: JP-196868.
PA (SUNR) SUNTORY LTD.
DR NPI: 97-314231/29.
DR Mortierella alpina cytochrome b5 (cytb5) gene - used for recombinant
DR production of cytb5 which is used in the preparation of essential
DR fatty acids
PS Claim 9: Pages 9-10: 14pp; Japanese.
CC The present sequence encodes the Mortierella alpina cytochrome b5
CC (cytb5); useful for the efficient preparation of human essential
CC fatty acids.
SQ Sequence 3420 BP; 820 A; 852 C; 858 G; 884 T;
Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 3420;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PN N09402599-A.
PD 03-FEB-1994.
PF 19-JUL-1993: U06754.
PR 22-JUL-1992: US-525985.
PI (SUNTORY) SUNTORY CO., LTD.
PI Bunkyo, Tokyo, Japan.
DR NPI: 94-048857/06.
DR Advanced glycosylation end-products, typically in the form of
DR age-plasmids, can be transfected into cells and used to capture
DR transposons.
PS Claim 6: Fig 4: 5pp; English.
CC The DNA sequence comprises a portion of a transposon INS-8 that
CC affects expression and related cellular activity. The DNA has been
CC linked with advanced glycosylation end products and is typically in
CC the form of a plasmid. The plasmid may be used to capture transposons
CC AGE modification of the plasmid may activate the transposons which
CC are captured. Such capture or movement of transposons in a cell may
CC be used to treat tumour cells.
CC See also Q57072-A.
SQ Sequence 3425 BP; 975 A; 767 C; 700 G; 981 T;
Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 3425;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2627 tcaacacaaatcat 2640
Cp 527 TCACCAAAATCAT 540

RESULT 436
ID T96832 standard; DNA: 3507 BP.
DT 27-SEP-1998 (first entry)
DE Intron 2 of human p53/p130 tumour suppressor gene.
KW Retinoblastoma tumour suppressor gene; p53; p130; p53/p130 gene;
KW cell cycle; tumour suppressor gene; cancer; molecular marker;
KW diagnosis; prognosis; predisposition; endometrial carcinoma;
KW human; intron 2; ss.
OS Homo sapiens.
PR J09738125-A1.
PR 03-APR-1997: U05598.
PR 03-MAR-1997: US-035532.
PR 05-APR-1996: US-014943.
PR 05-JUN-1996: US-013372.
PR 05-JUN-1996: US-013372.
PA (UVA) UNIV. JEFFERSON THOMAS.
PI Baldi A, Giordano A.
DR NPI: 97-512731/47.

PT Tumour suppressor pRb2/p130 gene intron and promoter sequences -
 PA used for the diagnosis and prognosis of cancer and predicting
 PT Clin 41, Page 121-123; 169pp; English.
 CC This genomic DNA sequence comprises intron 2 of the human
 CC tumour suppressor pRb2/p130 gene that plays a key role in cell
 CC cycle progression. The gene was isolated from a human P1 genomic
 CC library. The expression level of the pRb2/p130 gene in a
 CC sample from a patient can be determined and used to: (i) determine
 CC cancerous disease status; (ii) detect a
 CC risk for cancer; or (iii) grade a cancer; where the level of
 CC expression of pRb2/p130 is indicative of the grade of the cancer,
 CC prognosis or the presence of cancer. The sequence is an unfavourable
 CC provides methods for identifying polymorphisms and mutations in an
 CC exon of the pRb2/p130 gene, especially using primers (see
 CC T96833-96) based on the 3' noncoding region. The promoter (see
 CC T96833-96) can be used for the diagnosis and prognosis of cancer and for
 CC prediction of predisposition to cancer, particularly endometrial
 CC carcinoma, ovarian cancer, a squamous cell carcinoma of the lung,
 CC or adenocarcinoma of the lung.
 CC Sequence 3507 BP; 1051 A; 619 C; 725 G; 1112 T;
 CC
 CC Query Match 1.34; Score 14; DB 38; Length 3507;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 540 ttgctgcttag 553
 CC |||||||||||
 CC Cp 726 ttgctgcttag 713

RESULT 437

ID Q86693 standard; DNA: 3508 BP.
 AC 288359.1995 (first entry)
 DE Candida tropicalis pR233 isocitrate lyase.
 DE Candida tropicalis pR233; isocitrate lyase; ICL; promoter;
 KW eukaryotic; prokaryotic; expression controlling DNA; ds.
 PW Candida tropicalis.
 PS Location/Qualifiers
 FT cds 1531..3183
 FT /tag= a
 FT 307059576-A.
 PW 307059576-A.
 PS 307059576-A.
 PI 24-AUG-1993; 209705.

PR 24-AUG-1993; JP-209705.
 PA (MITO) MITSUBISHI KASEI CORP.
 DR WPI: 95-153900/18.
 PT An expression-controlling DNA and an expression vector contg. the
 PT DNA "useful for the preparation of heterologous proteins
 CC Example 1; Pages 7-9; 16pp; Japanese.
 CC Q86693 encodes 872576 Candida tropicalis pR233 isocitrate lyase
 CC protein. The protein is a 100 kDa protein. The protein is a
 CC the claimed promoter sequence Q86694-Q86705 were prep'd. These
 CC promoter sequences can be used in the prep'n. of heterologous
 CC proteins, in eukaryotic and prokaryotic host cells.
 CC Sequence 3508 BP; 1053 A; 752 C; 781 G; 922 T;
 CC
 CC Query Match 1.34; Score 14; DB 15; Length 3508;
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 508 cctactcttag 521
 CC |||||||||||
 CC Cq 128 cctactcttag 141

RESULT 438

ID T31259 standard; cDNA: 3511 BP.
 AC T31259.1996 (first entry)
 DE PM-1A-RAR-alpha DNA.
 DE PM-1A-RAR-alpha DNA.
 KW lymphoma; therapy; PM-1A-RAR-alpha; retinoic acid receptor;
 KW external guide sequence; EGS; antisense; ss.
 PS Homo sapiens.
 PS Location/Qualifiers
 FT misc_feature 1..1724
 FT /tag= a
 FT misc_difference 1725..3511
 FT /tag= b
 FT /note= "base n at position 1477 is not identified
 FT in the specification"
 FT misc_feature 1725..3511
 FT /tag= c
 FT /note= "RAR-alpha DNA"
 PW09618733-A2.
 PD 20-JUN-1996.
 PD 14-DEC-1994; US-334356.
 PD (INNOVIR) INNOVIR LAB INC.
 PI George ST, Goldberg AR, Pace U;
 DR WPI: 96-300650/30.
 PT RNA construct(s) including ribosome(s) and antisense

PT oligo-nucleotide(s) - for the inactivation of RNA associated with,
 CC e.g. promyelocytic leukemia or follicular lymphoma
 CC Example 1; Page 50-52; 81pp; English.
 CC PM-1A-RAR-alpha DNA is associated with acute promyelocytic
 CC leukemia (APL). APL is a type of leukemia that occurs in the local
 CC translocation between the long arms of chromosomes 15 and 17
 CC resulting in a fusion of the retinoic acid receptor gene (RAR-alpha)
 CC and a gene for a putative transcription factor, PM1. Methods for
 CC and antisense oligonucleotides for the treatment of APL are
 CC specifically cleave the PM1-RAR-alpha fusion mRNA (see also T31251
 CC but not wild-type RAR-alpha mRNA). Plasmids carrying portions of the
 CC PM1-RAR-alpha and RAR-alpha genes (see also T31260) allow synthesis
 CC screening processes.
 CC Sequence 3511 BP; 715 A; 1168 C; 1038 G; 599 T;
 CC
 CC Query Match 1.34; Score 15; DB 22; Length 3511;
 CC Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 1824 acagctcttag 1838
 CC QY 995 acagctcttag 1009

RESULT 439

ID T31259 standard; RNA: 3511 BP.
 AC T31259.1996 (first entry)
 DE PM-1A-RAR-alpha RNA sequence.
 DE PM-1A-RAR-alpha RNA sequence.
 KW lymphoma; therapy; PM-1A-RAR-alpha; retinoic acid receptor;
 KW external guide sequence; EGS; antisense; ss.
 PS Homo sapiens.
 PS Location/Qualifiers
 FT Key
 FT misc_feature 1725..3511
 FT /tag= a
 FT misc_difference 1477
 FT /tag= b
 FT /note= "base n at position 1477 is not identified
 FT in the specification"
 FT misc_feature 1725..3511
 FT /tag= c
 FT /note= "RAR-alpha RNA"
 FT misc_feature 1725..3511
 FT /tag= d
 FT /note= "cleavage site 1"
 FT misc_feature 1747..1749
 FT /tag= e

PT W09618733-A2.
 PD 20-JUN-1996.
 PD 14-DEC-1995; US-334356.
 PD (INNOVIR) INNOVIR LAB INC.
 PI George ST, Goldberg AR, Pace U;
 DR WPI: 96-300650/30.
 CC RNA construct(s) including ribosome(s) and antisense
 CC oligonucleotide(s) for the inactivation of RNA associated with,
 CC e.g. promyelocytic leukemia or follicular lymphoma
 CC Example 1; Page 53-55; 81pp; English.
 CC PM-1A-RAR-alpha RNA (T31246) is associated with acute promyelocytic
 CC leukemia (APL). APL is characterized by a balanced and
 CC translocation between the long arms of chromosomes 15
 CC and 17, resulting in a fusion of the retinoic acid receptor gene
 CC (RAR-alpha, see also T31260) and a gene for a putative
 CC transcription factor, PM1. Methods for treating APL involve the
 CC use of ribosomes, external guide sequences, and antisense
 CC oligonucleotides. The PM1-RAR-alpha fusion mRNA is specifically
 CC cleaved by the PM1-RAR-alpha fusion mRNA (see also T31251) but not
 CC wild-type RAR-alpha mRNA.
 CC Sequence 3511 BP; 713 A; 1168 C; 1038 G; 591 U;
 CC
 CC Query Match 1.34; Score 15; DB 22; Length 3511;
 CC Best Local Similarity 80.0%; Pred. No. 8.37e+01;
 CC Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 1824 acagctcttag 1838
 CC QY 995 acagctcttag 1009

RESULT 440

ID T80106 standard; cDNA: 3526 BP.
 AC T80106.
 DT 28-NOV-1997 (first entry)
 DE Human KP-1 gene clone KP1-1.
 DE Human KP-1 gene clone KP1-1.
 KW Albinism; disease; KP-1; human; mouse; mouse; antibody; ss.
 PS Homo sapiens.
 PS J05215495-A.
 PD 19-AUG-1997.
 PD 19-AUG-1997.
 PD 07-DEC-1995; JP-345559.
 PA (SUMO) SUMOMO SEITAKU KK.
 DR WPI: 97-465073/43.
 CC A brain-specific expression gene - used in the diagnosis of
 CC brain-specific expression gene - used in the diagnosis of
 CC Example 3; Page 14-15; 23pp; Japanese.
 CC T80106 and T80107 represent clones of the human KP-1 gene (see T80099 for
 CC wild type sequence). This sequence, T80098, and T80100 all represent cDNA

Db 2113 atatactctacg 2126
 Cp 27 ATATATCTTCACTG 14

RESULT 450
 ID T60045 standard: DNA: 3738 BP.
 AC T60045:
 DT 14-MAY-1997 (first entry)
 DE 86Q3a toxin coding sequence
 EE Bacillus thuringiensis, hymenopterans pest, pharaoth ant;
 KW biological control, Monomorium pharaonis; delta-endotoxin, lepidoptera;
 KW insect; da.
 OS Bacillus thuringiensis isolate PS86Q3.
 PN US3596072 A
 PP 22-MAY-1991: 703977.
 PR 22-MAY-1991: US-703977.
 PR 25-NOV-1991: US-797645.
 PR 24-NOV-1991: US-156232.
 PA (MCO) MYCOGEN CORP.
 PI Foncestrada L, Fu J, Kennedy MK, Meier H, Payne JM,
 PI Randall JB, Schnef HP, Schwab GE, Ulick RU;
 DR P-PSDB: w1872.
 DT Bacillus thuringiensis toxin - active against hymenopterans pests
 PS Disclosure: Column 63-68; 64pp; English.
 CC This sequence represents the coding sequence for the 86Q3a toxin isolated
 CC from a strain of *Bacillus thuringiensis* (strain 86Q3a) characterized by
 CC its ability to form a gran-positive spore forming, soil bacterium, characterized by
 CC its parasitoid crystalline protein inclusions. These proteins can be highly
 CC toxic to pests, and have been used to produce insect resistant plants.
 CC The previously isolated h.t. delta-endotoxins were mainly active against
 CC lepidopteran insects. This is the first time that a delta-endotoxin has been
 CC isolated from a bacterium that is active against hymenopterans. This is an
 CC example of a toxin of the invention, for which the sequences shown in
 CC w13888 and w13871 represent the generic formulae. As the toxins of the
 CC invention are active against hymenopterans pests, they can be used for the
 CC biological control of ants, particularly pharaoth ants (*Monomorium*
 CC *pharaonis*).
 SQ Sequence: 3738 BP: 1305 A; 602 C; 686 G; 1145 T;

Query Match 1.38; Score 14; DB 29; Length 3738;
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2113 atatactctacg 2126
 Cp 27 ATATATCTTCACTG 14

RESULT 451
 ID Q31412 standard: DNA: 3738 BP.
 AC Q31412:
 DT 20-APR-1993 (first entry)
 DE Bacillus thuringiensis gene 86Q3(a).
 EE Bacillus thuringiensis isolate PS86Q3.
 OS Bacillus thuringiensis isolate PS86Q3.
 PN 1.3738
 PP 1.3738
 PR 26-NOV-1992.
 PR 22-MAY-1992: 004116.
 PR 25-NOV-1991: US-797645.
 PR 25-NOV-1991: US-797645.
 PA (MCO) MYCOGEN CORP.
 PI Kennedy MK, Meier H, Payne JM, Randall JB, Ulick RU;
 DR P-PSDB: w1872.
 DT Toxin proteins isolated from Bacillus thuringiensis - for controlling
 PT ants. e.g. fire, carpenter, Argentine and pharaoth ants
 PS Claim 15, Page 48; 71pp; English.
 CC This sequence represents the coding sequence for the toxin protein
 CC 86Q3(a). This is useful as a method of biological control of ants,
 CC e.g. fire ants, carpenter ants, Argentine ants and pharaoth ants,
 CC as an alternative to chemical insecticides.
 SQ Sequence: 3738 BP: 1305 A; 602 C; 686 G; 1145 T;

Query Match 1.38; Score 14; DB 5; Length 3738;
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2113 atatactctacg 2126
 Cp 27 ATATATCTTCACTG 14

FT signal_peptide 152..211
 FT mat_peptide 212..3520
 FT 706315382-A
 PD 15-NOV-1994.
 PR 06-MAY-1993: 129912.
 PR 06-MAY-1993: JP-129912.
 PA (SUDA) SUDA PHARM CO LTD.
 DR WPI: 95-03231/03.
 DR P-PSDB: 867391.
 DT DNA coding a tie-2 receptor and a tie-2 receptor - used in the
 PT diagnosing Hematolymphatic diseases.
 PT The cDNA encodes R67391, the murine tyrosine kinase receptor tie-2.
 CC The cDNA is thought to participate in hematolymphatic, and can therefore
 CC be used in the diagnosis of hematolymphatic diseases.
 SQ Sequence: 3760 BP: 1024 A; 876 C; 970 G; 890 T;

Query Match 1.38; Score 14; DB 13; Length 3760;
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2833 agcgtattttac 2846
 QY 108 AGCCTATTCTTACC 121

RESULT 453
 ID T61864 standard: DNA: 3780 BP.
 AC T61864:
 DT 06-JUN-1997 (first entry)
 DE c-Abl gene; antisense RNA; cancer; therapy; DNA damage;
 KW radiotherapy; mitomycin C; cytostatic; tyrosine kinase; as.
 OS Homo sapiens.
 PN 069708184 A1.
 PP 30-AUG-1995: US-520923.
 PR 30-AUG-1995: US-520923.
 PA (ARCH-) ARCH DEV CORP.
 PI LOMAD DNA FARMER CANCER INST INC.
 DR WPI: 97-179167/16.
 DT Antisense RNA molecule that inhibits expression of c-Abl gene
 PT product - useful in killing malignant cells in cancer, optionally in
 PT combination with DNA damaging agents.
 CC The c-Abl gene (761864 and 761865) encodes a non-receptor tyrosine
 CC kinase that is activated DNA damaging agents. Antisense constructs

CC directed to the c-abl gene can be used to selectively inhibit the
 CC expression c-Abl following exposure of cells to a DNA damaging
 CC agent such as ionising radiation or mitomycin C. When used in
 CC combination with DNA damaging agents, the antisense RNA provides a
 CC synergistic effect, killing effect useful for killing malignant cells
 CC in cancer treatment.
 SQ Sequence: 3780 BP: 853 A; 1136 C; 1166 G; 625 T;

Query Match 1.38; Score 14; DB 29; Length 3780;
 Best Local Similarity 100.0%; Pred. No. 4.47e-02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2737 ggaagctctgag 2750
 QY 602 GGAAGCTCTGAG 615

RESULT 454
 ID T95244 standard: cDNA: 3804 BP.
 AC T95244:
 DT 21-MAY-1998 (first entry)
 DE Human hematopoietic stem and progenitor cell antigen AC133 cDNA.
 KW AC133: hematopoietic stem cell; progenitor cell; antigen;
 OS marker; human; bone marrow; gene therapy; da.
 PN 97-543778/50.
 PP 38..2635
 PR 26-APR-1996; US-639891.
 PA (AMCE-) AMCELL CORP.
 PI Miraglia S, Yin A;
 DR P-PSDB: w26769.
 DT Anticbody specific for hematopoietic stem/progenitor cell antigen.
 PT AC133 - useful to purify such cells, e.g. for use as gene therapy
 PT reagent for bone marrow transplantation.
 PS Claim 21, Page 42-48; 78pp; English.
 CC This claimed nucleic acid molecule includes a coding region for
 CC a human hematopoietic progenitor cell antigen (see w26769),
 CC designated AC133 antigen. It was isolated from a WMR-Rb-1
 CC cell line, which is a human hematopoietic stem cell line. The
 CC cDNA is highly tissue specific. It is detected on a subset of
 CC

PS Disclosure: Fig 13, 11pp; Japanese.
 CC cloned using probes derived from animal amidating enzyme has been.
 CC The C-terminal amidating enzyme can be used to amidate.
 CC Physiological active peptides such as calcitonin and gastrin.
 CC It also amides 1-tyr-tyr-gly, phe-gly-phe-gly and gly-phe-gly.
 CC Sequence 3225 BP; 1146 A; 789 C; 899 G; 1091 T;
 SQ Sequence 3225 BP; 1146 A; 789 C; 899 G; 1091 T;
 Query Match 1.34; Score 15; DB 13; Length 3925;
 Best Local Similarity 100.04; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3708 tttcttcacccaaa 3722
 QY 522 tttcttcacccaaa 536
 RESULT 468
 ID Q03736 standard; cDNA: 3928 BP.
 AC 16-AUG-1990 (first entry)
 DE Sequence of plasmid pCER11 encoding cDNA of fowl preprolactin.
 DT 16-AUG-1990 (first entry)
 PR 16-AUG-1990: JP-203913.
 PA (NIGP-) Nippon Gene KK.
 DR P-PSB8-203913/14.
 PT used for avian breeding and proliferation.
 PT Recombinant avian prolactin and recombinant avian preprolactin -
 PS disclosure: Fig 1A-G; Pp; Japanese.
 CC The mature prolactin derived was used to induce maternal behaviour in
 CC the mouse prolactin derived was used to induce maternal behaviour in
 CC fowl and encourage breeding and proliferation.
 CC Sequence 3926 BP; 1047 A; 966 C; 938 G; 975 T;
 SQ Sequence 3926 BP; 1047 A; 966 C; 938 G; 975 T;
 Query Match 1.34; Score 15; DB 13; Length 3926;
 Best Local Similarity 100.04; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1036 ctttcgtttttata 1040
 QY 183 ctttcgtttttata 197
 RESULT 469
 ID 791902 standard; cDNA: 4000 BP.
 AC 07-MAR-1998 (first entry)
 DE Mannose-1-phosphate transferase gene MNM4
 DT 07-MAR-1998 (first entry)
 PR 29-MAR-1998: JP-075667.
 PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 DR WP1: 97-553460/51.
 PT useful for high mannose type neutral saccharide chain production
 PT Positive; 430763 story gene of mannose-1-phosphate transfer in yeast -
 CC This sequence represents the gene of the invention, designated MNM4,
 CC encoding a protein which positively regulates mannose-1-phosphate
 CC transfer in a yeast. The gene is useful for the preparation of human high
 CC mannose type neutral saccharide chain.
 CC Sequence 4000 BP; 1453 A; 750 C; 770 G; 1027 T;
 SQ Sequence 4000 BP; 1453 A; 750 C; 770 G; 1027 T;
 Query Match 1.34; Score 15; DB 38; Length 4000;
 Best Local Similarity 100.04; Pred. No. 8.37e+01; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2013 ttttcgtttttata 2027
 QY 33 ttttcgtttttata 47
 RESULT 470
 ID Q84303 standard; cDNA: 4010 BP.
 AC Q84303;
 DE Human NF-AT120, X subfamily gene.
 DT 12-AUG-1995 (first entry)
 PR Human NF-AT120, X subfamily gene.
 DE Human NF-AT120, nuclear factor of activated T cells 120 protein;
 KW transcription; differentiation; cytokine; T-cell; T-lymphocyte; ss.

OS Homo sapiens.
 PH Key Location/Qualifiers
 FH cds 304..3531
 FT cds 1756..1758
 FT unsure
 FT /tag= b
 FT /note= "bases 1756-1758 may be replaced by gta,
 FT gaa or gga"
 FT /tag= c
 FT /note= "bases 3090-3092 may be replaced by agt,
 FT aga or agg"
 FT W0502053-A
 PR 05-JUL-1994: U07297.
 PR 06-JUL-1993: US-088483.
 PR 30-JUL-1993: US-099998.
 PR 10-AUG-1993: US-113971.
 PR 05-NOV-1993: US-148061.
 PR 04-APR-1994: US-222626.
 PA (SCHE) SCHERING CORP.
 PI Arai N, Masuda ES, Tokumitsu H;
 DR WP108-06699/09.
 PT Nuclear factor of Activated T cells 120 protein, antibody and
 PT nuclear acid - useful as transcriptional regulator of genes, such
 PT as cytokines.
 CC disclosure: Page 78-83, 10pp; English.
 CC given in Q84300 and encoded NF-AT120 (Nuclear Factor of
 CC Activated T cells protein 120) (R66873). Amplified DNA was used
 CC to screen for related NF-AT proteins by hybridization, leading to
 CC the discovery of 3 related subfamilies, designated class C, P and
 CC X (R66874, R66875, R66876, R66877, R66878, R66879, R66880, R66881,
 CC X (R66874, R66875, R66876, R66877, R66878, R66879, R66880, R66881,
 SQ Sequence 4010 BP; 1026 A; 1013 C; 902 G; 1069 T;
 Query Match 1.34; Score 14; DB 13; Length 4010;
 Best Local Similarity 100.04; Pred. No. 3.47e+02; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3743 ctttcgtttttata 3756
 QY 183 ctttcgtttttata 196
 RESULT 471
 ID Q84303 standard; cDNA: 4010 BP.
 AC Q84303;
 DE Human NF-AT120, X subfamily gene.
 DT 12-AUG-1995 (first entry)
 PR Human NF-AT120, X subfamily gene.
 DE Human NF-AT120, nuclear factor of activated T cells 120 protein;

transcription; differentiation; cytokine; T-cell; T-lymphocyte; ss.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FH cds 304..3531
 FT cds 1756..1758
 FT unsure
 FT /tag= b
 FT /note= "bases 1756-1758 may be replaced by gta,
 FT gaa or gga"
 FT /tag= c
 FT /note= "bases 3090-3092 may be replaced by agt,
 FT aga or agg"
 FT W0502053-A
 PR 05-JUL-1994: U07297.
 PR 06-JUL-1993: US-088483.
 PR 30-JUL-1993: US-099998.
 PR 10-AUG-1993: US-113971.
 PR 05-NOV-1993: US-148061.
 PR 04-APR-1994: US-222626.
 PA (SCHE) SCHERING CORP.
 PI Arai N, Masuda ES, Tokumitsu H;
 DR WP108-06699/09.
 PT Nuclear factor of Activated T cells 120 protein, antibody and
 PT nuclear acid - useful as transcriptional regulator of genes, such
 PT as cytokines.
 CC disclosure: Page 78-83, 10pp; English.
 CC given in Q84300 and encoded NF-AT120 (Nuclear Factor of
 CC Activated T cells protein 120) (R66873). Amplified DNA was used
 CC to screen for related NF-AT proteins by hybridization, leading to
 CC the discovery of 3 related subfamilies, designated class C, P and
 CC X (R66874, R66875, R66876, R66877, R66878, R66879, R66880, R66881,
 SQ Sequence 4010 BP; 1026 A; 1013 C; 902 G; 1069 T;
 Query Match 1.34; Score 14; DB 13; Length 4010;
 Best Local Similarity 100.04; Pred. No. 3.47e+02; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3311 ttttcgtttttata 3324
 QY 605 ttttcgtttttata 592
 RESULT 472
 ID Q84303 standard; cDNA: 4010 BP.
 AC Q84303;
 DE Human NF-AT120, X subfamily gene.
 DT 12-AUG-1995 (first entry)
 PR Human NF-AT120, X subfamily gene.
 DE Human NF-AT120, nuclear factor of activated T cells 120 protein;

US-08-887-977-9.FMD

```

FT      /tag= C deleted in Q20199 and Q20200*
PD      J03262484-A.
PE      22-NOV-1991.    051905
PF      14-MAR-1992.    J0563106.
PG      (SHIS ) SHISEIDO KK.
PH      WPI : 92-012701/02.
PI      P-FDSB; R300S?   coding for peptide C terminal amidation enzyme
PJ      or'd. from horse
PK      Claim 1, Fig 1; 18pp; Japanese.
PL      The sequence was determined from a clone isolated from a library
PM      constructed from cDNA from horse stomach muscle. The sequence
PN      is one of four amino acids (Glu, Thr, Ser, Lys) which cause the
PO      sequence up to base 7417 where they diverge, encoding different
PP      C-terminal. The different termini are created by deletions in the
PQ      C-terminal. One section of DNA has been sequenced (see below).
PR      The sequence encodes an amidase (see Q20199, Q20200).
PS      Similarity 4045 BP; 1181 A.; 825 G; 922 G; 1117 T;
PT      Query Match 1 3%; Score 15; DS 3; Length 4045;
PU      Identical 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PV      Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PW      DB 3828 tgccttcacaccacaa 3942
PX      522 tgccttcacaccacaa 536
PY      522 tgccttcacaccacaa 536

```

KW	renal; hepatoma; stomach; cancer; tumour rejection antigen precursor;
KN	RAB; disease; diagnosis; antibody; cytotoxic T cell; ds.
FT	HMO asplasm.
CS	Key
PD	Location/Qualifiers
FT	35..2566
FT	/tag= a
FT	/product= brain glycogen phosphorylase
FT	W9748817.A1
PD	31-OCT-1997
PP	25-JUN-1997; U11089.
PP	25-JUN-1996; UR-672351.
PA	(LIDM.) LIDMG INST CANCER RES.
PA	REGG UNIV CALIFORNIA,
PT	SACRAMENTO, CA 95834-1201
DR	WS1: 98-07180/07
DR	WS2: 98-07180/07
DR	WS3: 98-07180/07
DR	WS4: 98-07180/07
DR	WS5: 98-07180/07
DR	WS6: 98-07180/07
DR	WS7: 98-07180/07
DR	WS8: 98-07180/07
DR	WS9: 98-07180/07
DR	WS10: 98-07180/07
DR	WS11: 98-07180/07
DR	WS12: 98-07180/07
DR	WS13: 98-07180/07
DR	WS14: 98-07180/07
DR	WS15: 98-07180/07
DR	WS16: 98-07180/07
DR	WS17: 98-07180/07
DR	WS18: 98-07180/07
DR	WS19: 98-07180/07
DR	WS20: 98-07180/07
DR	WS21: 98-07180/07
DR	WS22: 98-07180/07
DR	WS23: 98-07180/07
DR	WS24: 98-07180/07
DR	WS25: 98-07180/07
DR	WS26: 98-07180/07
DR	WS27: 98-07180/07
DR	WS28: 98-07180/07
DR	WS29: 98-07180/07
DR	WS30: 98-07180/07
DR	WS31: 98-07180/07
DR	WS32: 98-07180/07
DR	WS33: 98-07180/07
DR	WS34: 98-07180/07
DR	WS35: 98-07180/07
DR	WS36: 98-07180/07
DR	WS37: 98-07180/07
DR	WS38: 98-07180/07
DR	WS39: 98-07180/07
DR	WS40: 98-07180/07
DR	WS41: 98-07180/07
DR	WS42: 98-07180/07
DR	WS43: 98-07180/07
DR	WS44: 98-07180/07
DR	WS45: 98-07180/07
DR	WS46: 98-07180/07
DR	WS47: 98-07180/07
DR	WS48: 98-07180/07
DR	WS49: 98-07180/07
DR	WS50: 98-07180/07
DR	WS51: 98-07180/07
DR	WS52: 98-07180/07
DR	WS53: 98-07180/07
DR	WS54: 98-07180/07
DR	WS55: 98-07180/07
DR	WS56: 98-07180/07
DR	WS57: 98-07180/07
DR	WS58: 98-07180/07
DR	WS59: 98-07180/07
DR	WS60: 98-07180/07
DR	WS61: 98-07180/07
DR	WS62: 98-07180/07
DR	WS63: 98-07180/07
DR	WS64: 98-07180/07
DR	WS65: 98-07180/07
DR	WS66: 98-07180/07
DR	WS67: 98-07180/07
DR	WS68: 98-07180/07
DR	WS69: 98-07180/07
DR	WS70: 98-07180/07
DR	WS71: 98-07180/07
DR	WS72: 98-07180/07
DR	WS73: 98-07180/07
DR	WS74: 98-07180/07
DR	WS75: 98-07180/07
DR	WS76: 98-07180/07
DR	WS77: 98-07180/07
DR	WS78: 98-07180/07
DR	WS79: 98-07180/07
DR	WS80: 98-07180/07
DR	WS81: 98-07180/07
DR	WS82: 98-07180/07
DR	WS83: 98-07180/07
DR	WS84: 98-07180/07
DR	WS85: 98-07180/07
DR	WS86: 98-07180/07
DR	WS87: 98-07180/07
DR	WS88: 98-07180/07
DR	WS89: 98-07180/07
DR	WS90: 98-07180/07
DR	WS91: 98-07180/07
DR	WS92: 98-07180/07
DR	WS93: 98-07180/07
DR	WS94: 98-07180/07
DR	WS95: 98-07180/07
DR	WS96: 98-07180/07
DR	WS97: 98-07180/07
DR	WS98: 98-07180/07
DR	WS99: 98-07180/07
DR	WS100: 98-07180/07

Page 469

```

FT      /+tag= B  

FF      2602..2606  

FT      /+tag= h  

FF      3031..3046  

exon    /+tag= i  

PD      W09120237-A.  

PD      14-OCT-1993.   U01203.  

FF      01-APR-1993;   U01203.  

FF      01-APR-1993;   U01203.  

FA      (CAB) CAMBRIDGE BIOTECHNOLOGY INC.  

PI      Johnson CD, Marchionni MA;  

DI      Wfir_93_33694742.  

PR      Long-distance homology cloning of genes from lower organisms -  

PT      used to identify DNA that codes for evolutionary conserved  

PT      aminoacid sequences 188pp. English  

CC      The primers/probes (04596G-04325S), are used to isolate the cnd-4  

CC      gene from the nematode C. briggsae.  

CC      Sequence 4093 BP; 1226 A; 792 C; 725 G; 1346 T;  

Query Match 1.3%; Score 14; DB S; Length 4093;  

Beat Local Similarity 100.0%; Pred. No. 3.47e+02;  

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0; O: Gaps 0;  

DB      1937 tgacaccacagca 1950  

CP      |||||  

CP      181 tgATACCAcAcaga 168

```

RESULT 479

ID Q4062 standard; DWA, 4100 BP.

DT 28-JAN-1996 (first entry)

DE The rat beta-actin gene.

KW Beta-actin; primer: mRNA; specificity: pharmaceutical; as.

PR J07123984-A.

PD 16-MAY-1995.

PP 05-NOV-1995; 275832.

FA (SIB) HITACHI CHUKO CO LTD.

WP1: 95-311667/28.

WP2: 95-311667/28.

A primer for the detection and the determ. of a specific messenger RNA which detect and determine specific mRNA(s) with high sensitivity.

Example 277 Page 28-39; Japanese.

Q4062 is the rat beta-actin gene. This gene is amplified by the primers Q74063 and Q74064. The primers are used specifically for the PCR amplification of the rat beta-actin gene. The advantage of the primers Q74063 and Q74064 is that they have a high percentage of GC base pairs, high sensitivity and reliability, and are useful for the detection of the rat beta-actin gene.

CC Industry. 4100 BP; 765 A; 1111 C; 1136 G; 1097 T;
 SQ Sequence 4100 BP; 765 A; 1111 C; 1136 G; 1097 T;
 Query Match 1.38; Score 14; DB 16; Length 4100;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1798 ttcagcagcagtg 1811
 |||||||
 QY 1104 AGCTGAGTCTCCCT 1117

RESULT 480
 ID T58545 standard; cDNA; 4134 BP.
 AC T58545;
 DT 25-MAR-1997 (first entry) - kinase PI3K-gamma cDNA.
 KW Phosphatidylinositol 3-kinase; PI3K; gamma; immunosay;
 KW histamine secretion; nerve cell differentiation; glucose transport;
 KW modulation; regulation; Alzheimer's disease; lipolysis; ds.
 Key words sapiens. Location/Qualifiers
 cds
 423..3572
 /*tag= a
 /*product= PI3K-gamma
 DB4445562-C1.
 PD 04-APR-1996.
 PF 20-DEC-1994; 445562.
 PR 13-OCT-1994; DE-436696.
 PT PLAC 7 MAX PLANKER GES FORDERUNG WISSENSCHAFTEN.
 PI WPI: 96-172545/18.
 DR P-PDB: W1576.
 DT New phosphatidylinositol 3-kinase protein - useful as immunogen and
 PT for determining of kinase activity
 CC A 402 bp cDNA fragment was amplified from a human bone marrow
 CC library using PCR primers corresponding to amino acid sequences
 CC KMDPDR and HDPG. The amplified fragment was used to probe a human
 CC cDNA library and several overlapping clones were isolated.
 CC The cDNA clones were sequenced and the protein was expressed in
 CC of 1049 residues. The protein is a novel phosphatidylinositol 3-
 CC kinase (PI3K) that differs in its regulatory mechanism from the
 CC known PI3K-alpha and -beta enzymes. The new enzyme has been
 CC subcloned and used as an immunogen. The enzyme,
 CC antibodies against it, and the cDNA fragment were used as an immunogen.
 CC modulating cell proliferation, receptor-mediated signal transduction,
 CC histamine secretion, nerve cell differentiation, glucose transport,
 CC and anti-lipolytic activity or for treating Alzheimer's disease.
 CC N.B. Although the claimed sequences are referred to by SEQ ID.

CC Numbers. A sequence listing did not appear in the original printed
 CC patent application.
 SQ Sequence 4134 BP; 1127 A; 999 C; 999 G; 1010 T;
 Query Match 1.38; Score 14; DB 26; Length 4134;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3079 ttcagcagcagtg 3092
 |||||||
 QY 908 TTCAGCAGCAGTG 895

RESULT 481
 ID T58545 standard; cDNA; 4137 BP.
 AC T58545;
 DT 25-MAR-1997 (first entry) - kinase PI3K-gamma cDNA.
 KW Phosphatidylinositol 3-kinase; PI3K; gamma; immunosay;
 KW histamine secretion; nerve cell differentiation; glucose transport;
 KW modulation; regulation; Alzheimer's disease; lipolysis; ds.
 Key words sapiens. Location/Qualifiers
 cds
 423..3575
 /*tag= a
 /*product= PI3K-gamma
 DB4445562-C1.
 PD 04-APR-1996.
 PF 20-DEC-1994; 445562.
 PR 13-OCT-1994; DE-436696.
 PT PLAC 7 MAX PLANKER GES FORDERUNG WISSENSCHAFTEN.
 PI WPI: 96-172545/18.
 DR P-PDB: W1577.
 DT New phosphatidylinositol 3-kinase protein - useful as immunogen and
 PT for determining of kinase activity
 CC A 402 bp cDNA fragment was amplified from a human bone marrow
 CC library using PCR primers corresponding to amino acid sequences
 CC KMDPDR and HDPG. The amplified fragment was used to probe a human
 CC cDNA library and several overlapping clones were isolated.
 CC The cDNA clones were sequenced and the protein was expressed in
 CC of 1049 residues. The protein is a novel phosphatidylinositol 3-
 CC kinase (PI3K) that differs in its regulatory mechanism from the
 CC known PI3K-alpha and -beta enzymes. The new enzyme has been
 CC subcloned and used as an immunogen. The enzyme,
 CC antibodies against it, and the cDNA fragment were used as an immunogen.
 CC modulating cell proliferation, receptor-mediated signal transduction,
 CC histamine secretion, nerve cell differentiation, glucose transport,
 CC and anti-lipolytic activity or for treating Alzheimer's disease.
 CC N.B. Although the claimed sequences are referred to by SEQ ID.

CC Glucose transport and anti-lipolytic activity or for treating
 CC Alzheimer's disease.
 CC N.B. Although the claimed sequences are referred to by SEQ ID.
 CC Numbers. A sequence listing did not appear in the original printed
 CC patent application.
 SQ Sequence 4137 BP; 1127 A; 999 C; 1001 G; 1010 T;
 Query Match 1.38; Score 14; DB 26; Length 4137;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3082 ttcagcagcagtg 3095
 |||||||
 QY 908 TTCAGCAGCAGTG 895

RESULT 482
 ID Q2159 standard; DNA; 4155 BP.
 AC Q2159;
 DT 20-APR-1993 (first entry)
 DE BT toxin 17a.
 KW nematode worms; nematocidal; nematocidal toxin; agriculture; plants;
 KW Scropi; Pests; Cry proteins.
 Key words thuringiensis. Location/Qualifiers
 cds
 1..4155
 /*tag= a
 /*product= BT toxin 17a.
 DB 09-DEC-1992.
 PF 01-MAY-1992; 303969.
 PR 03-MAY-1991; US-693018.
 PT 21-JAN-1992; US-810050.
 PI (MTCO) MYCOGEN CORP.
 DR Fonceirada L, Narva EE, Payne JM, Schnepf HE, Schwab GE;
 WPI: 92-48825/50.
 DT New Bacillus thuringiensis - useful for
 KW control of animal or plant parasites; deoxyribonucleic acid
 CC coding sequences; transformed hosts and transgenic plants
 CC Claim 1(b); Page 23; 57pp; English.
 CC This sequence encodes the Bacillus thuringiensis delta-endotoxin
 CC having the amino acid sequence: (Phe17a)17a. The delta-endotoxin
 CC lambda-domain is a protein of 17a amino acids. The delta-endotoxin
 CC digested with Sau3A and size fractionated by electrophoresis
 CC (9-23kb used). Packaged phage were plated on E. coli RW51 and
 CC (9-23kb used). Packaged phage were purified and digested with
 CC recombined. Positive plaques were purified and digested with
 CC phage DNA preparation. Phage DNA was digested with EcoRI and
 CC separated by electrophoresis and reprobed. Two positive
 CC bands were present, clones contg. 4.5 and 2.7 kb EcoRI fragments.

CC These were cloned into pBlac, an E. coli/MT shuttle vector made up
 CC of replication origins from pUC16 and pUC19. This was used to
 CC transform E. coli MG22, and this grown on IPTG and XGAL media.
 CC White colonies were selected, purified, and either pUC1627 contg.
 CC the 2 kb P817a toxin fragment (this sequence) or pUC1628 contg.
 CC the 2 kb P817a toxin fragment (this sequence) and sequenced.
 SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;
 Query Match 1.38; Score 14; DB 6; Length 4155;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 402 aactttacttcgag 415
 |||||||
 QY 355 MCTTACTTCGCG 368

RESULT 483
 ID Q2083 standard; DNA; 4155 BP.
 AC Q2083;
 DT 01-APR-1993 (first entry)
 DE Toxin 17a.
 KW Endotoxin; acarides; pest; Two Spotted Spider; mite;
 KW Phycophagus; ds.
 Key words thuringiensis. Location/Qualifiers
 cds
 1..4155
 /*tag= a
 /*product= BT toxin 17a.
 DB 12-NOV-1992.
 PF 30-APR-1992; 003546.
 PR 03-MAY-1991; US-693018.
 PT 21-JAN-1992; US-810050.
 PI (MTCO) MYCOGEN CORP.
 DR Bagley AL, Cannon RJC, Payne JM;
 WPI: 92-48825/50.
 DT New Bacillus thuringiensis isolates and toxins - used for
 KW controlling acarid pests of livestock, fowl, stored prods. and
 CC plants
 CC Claim 1(b); Page 25-27; 62pp; English.
 CC This sequence encodes a toxin which is active against acarides and
 CC is obtainable from B. thuringiensis isolates P817a, P817b, 3322,
 CC P852a1, P852b1, P852c1 and P852c2 are given in Q20803-07 and Q20820-21
 CC respectively. The toxin is a protein of 4155 amino acids. The toxin
 CC used against non-phycophagus mites such as acarid pests of livestock,
 CC fowl and stored prods. The genes can be cloned and used to
 CC transform other hosts, which can be used to control mites, or in
 CC the control of acarid pests of livestock, fowl, stored prods. and
 CC plants.
 SQ Sequence 4155 BP; 1386 A; 668 C; 819 G; 1280 T;
 Query Match 1.38; Score 14; DB 5; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 aacttactgag 415
QY 355 AACCTTACTGCG 368

RESULT 484

ID Q51687 standard; DNA; 4155 BP.
AC Q51687-1995 (first entry)
DE Bacillus thuringiensis crystal protein 17a gene.
KW Cystic protein: delta endotoxin; acariicide; pepticide;
KW Biological control agent; transgenic plant; crop improvement; da.
KW Bacillus thuringiensis; pharaoxin; delta-endotoxin; lepidoptera;
KW mat_peptide 1: 4155
FT mat_peptide 1: 4155
FT US262158-A.
PR 10-APR-1991: 693210.
PR 30-APR-1991: US-693210.
PR 13-SEP-1991: US-759248.
PR 30-SEP-1991: US-761841.
PR 10-SEP-1991: US-761841.
PR (MCO) MYCOGEN CORP.
PI Bagley AL, Cannon RJC, Payne JM;
DR WPI: 92-37787/47.
PR P-PSDB: R28803.
PR Bacillus thuringiensis delta-toxin cloned into microbe hosts.
PT Diclosure: Column 21-26; 42pp; English.
CC DNA encoding the insecticidal toxin can be cloned into baculo
CC virus vectors for expression in host microorganisms, preferably E. coli
CC strains which become resistant to the acariicide pests.
CC Specifically, the two-spotted spider mite (Tetranychus urticae) is
CC controlled. 4155 BP; 1386 A; 677 C; 811 G; 1281 T;
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 aacttactgag 415
QY 355 AACCTTACTGCG 368

RESULT 485

760069 standard; DNA; 4155 BP.
AC 760069-1997 (first entry)
DE 17a toxin coding sequence;
KW Toxin; ant; Bacillus thuringiensis; hymenopter pest; pharaox ant;
KW Biological control; Monomorium pharaonis; delta-endotoxin; lepidoptera;
KW Bacillus thuringiensis isolate PS17a.
PR 21-JAN-1997: 703977.
PR 22-MAY-1991: US-703977.
PR 25-NOV-1991: US-703977.
PR 22-MAY-1991: US-703977.
PR 24-NOV-1993: US-158232.
PR (MCO) MYCOGEN CORP.
PI Randall AL, Schnepf HE, Schwab GE, Ulick HJ,
DR WPI: 97-107615/10.
PR P-PSDB: W13884.
PR Bacillus thuringiensis toxin - active against hymenopter pests
PR This sequence represents the coding sequence for the 17a toxin isolated
CC from the Bacillus thuringiensis (B.t.) isolate PS17a. B.t. is a
CC gram-positive, spore forming, soil bacterium, characterised by parasporal
CC crystalline protein inclusions. These proteins can be highly toxic to
CC lepidopteran insects. The toxin is highly specific and is active against
CC previously isolated B.t. delta-endotoxins were mainly active against
CC lepidopteran insects, however the proteins of the invention are active
CC against hymenopter insects. The protein encoded by this sequence is an
CC insecticidal toxin of the invention, for which the sequences shown in
CC W13884 and W13717 are the preferred forms. The sequences shown in
CC the invention are active against hymenopter pests, they can be used for the
CC biological control of ants, particularly pharaox ants (Monomorium
CC pharaonis). 4155 BP; 1386 A; 677 C; 811 G; 1281 T;
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 aacttactgag 415
QY 355 AACCTTACTGCG 368

RESULT 486

ID Q10935 standard; DNA; 4155 BP.
AC Q10935-1993 (first entry)
DE 26-FEB-1996 (first entry)
DE 26-FEB-1996 (first entry)

nematode worms; nematocidal; nematocidal toxin; agriculture; plants;
KW crops; pests; Cryv proteins.
KW Bacillus Thuringiensis.
PR 10-APR-1991: 693210.
PR 30-APR-1991: US-693210.
PR 13-SEP-1991: US-759248.
PR 30-SEP-1991: US-761841.
PR 10-SEP-1991: US-761841.
PR (MCO) MYCOGEN CORP.
PI Bagley AL, Cannon RJC, Payne JM, Schnepf HE, Schwab GE;
DR WPI: 92-39886/48.
PR P-PSDB: R28803.
PR New genes and toxins against nematodes - obtd. from Bacillus
PR thuringiensis strains with nematocidal activity
PR This sequence encodes the Bacillus thuringiensis delta-endotoxin
CC having nematocidal activity. A library was constructed in
CC lambda vector using Bt strain 97 total cellular DNA partially
CC digested with EcoRI. The library was screened with a 1.5 kb
CC (9.2kb) used. Packaged phages were plated on E. coli NM522 and
CC screened, and positive plaques were purified and
CC phage DNA preparation. Phage DNA was digested with EcoRI and
CC bands were present on agarose gels. 4.5 and 2.7 kb EcoRI fragments
CC These were cloned into pUC19, an E. coli/ST shuttle vector made up
CC of replication origins from pUC19 and pUC19. This was used to
CC transform E. coli NM522, and this grown on IPTG and XGAL media.
CC The 2.7 kb PS17a toxin fragment (this sequence) or pMTG128 contg.
CC the 4.5 kb PS17a toxin fragment isolated and sequenced.
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 aacttactgag 415
QY 355 AACCTTACTGCG 368

RESULT 487

ID Q94053 standard; DNA; 4155 BP.
AC Q94053-1996 (first entry)
DE 26-FEB-1996 (first entry)
DE PS17a acariide-active toxin DNA from strain PS17.

Insecticide: toxin; acariide-active toxin; delta-endotoxin; PS17a; PS17.
KW two spotted spider mite; tetranychus urticae; bacillus thuringiensis; da.
KW Bacillus thuringiensis strain PS17.
PR 10-APR-1991: 693210.
PR 30-APR-1991: US-693210.
PR 13-SEP-1991: US-759248.
PR 30-SEP-1991: US-761841.
PR 10-SEP-1991: US-761841.
PR (MCO) MYCOGEN CORP.
PI Bagley AL, Cannon RJC, Payne JM;
DR WPI: 92-37787/47.
PR DNA encoding delta endotoxin of Bacillus thuringiensis - used for
PR the control of acariide pests and for the production of acariide pest
PR resistant plants
PR Example 3: Column 21-26; 42pp; English. Q94053 represent acariide-active
CC toxin genes isolated from B. thuringiensis isolate PS17 DNA. This
CC sequence represents PS17a. These sequences are delta-endotoxins.
CC Delta-endotoxins are insect toxins which act on insects in the orders
CC Lepidoptera and Diptera. The toxin acts to destroy insect gut epithelium
CC and disrupts the midgut. The toxin is highly specific and is active
CC transformed into biologically active forms by the insect gut juice
CC proteases (in susceptible insects). This sequence can be inserted into a
CC vector which can be introduced into plant cells, thereby conferring
CC insect resistance. The encoded protein can be combined with other
CC insect resistance genes to provide a complete control of the pest. This
CC complete control of target pests. Plants transformed with this sequence
CC become resistant to acariide pests, e.g. the two spotted spider mite,
CC Tetranychus urticae.
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 aacttactgag 415
QY 355 AACCTTACTGCG 368

RESULT 488

ID Q72179 standard; DNA; 4155 BP.
AC Q72179-1996 (first entry)
DE PS17a acariide-active toxin DNA from strain PS17.

16-WAN-1995 (first entry)
 Delta-WAN Thuringiensis PS17a delta-endotoxin gene.
 Endotoxin: Bacillus thuringiensis; PS17; toxin; delta-endotoxin.
 Parasitol: crystalline; inclusion; toxic; pest; spider mite;
 host: spider mite; pest; spider mite; pest; spider mite;
 phytophaga: mite; acarid; livestock; fowl; raptorial; insect;
 virus: plant; control; transgenic; resistance; da.
 OS Bacillus thuringiensis strain PS17.
 PP 11-SEP-1991: 759248.
 PP 13-SEP-1991: 759248.
 PP 13-SEP-1991: US-759248.
 PP 10-SEP-1991: US-768141.
 PP 10-SEP-1991: US-768141.
 PA (MTC) MYCOGEN CORP.
 PI Cannon JC, Payne J, Ralph AL;
 WI: 94-310547/38.
 PT Controlling house dust mite - using new delta endotoxin isolates
 from Bacillus thuringiensis with acaricidal activity
 Claim 5: Column 13-16: 10pp; English.
 PS The nucleotide sequence of a novel delta-type endotoxin from Bacillus
 thuringiensis strain PS17a (Genbank accession no. U000128) is
 1385 amino acids. The PS17 delta-endotoxin range in mol. wt. from 128
 kDa to 155 kDa. The delta-endotoxins are derived from parasporal crystallin
 protein inclusions and are highly toxic to pests and specific in their
 action. The delta-endotoxins are specific to the dust mite
 two-spotted spider mite (Tetranychus bioretii) and the dust mite
 (Dermatophagoides pteronyssinus). The use of B. thuringiensis containing
 the endotoxin can also target non-phytophagous mites such as acarid
 insects of which cockroach and stored products. The gene(s) of the B. th.
 isolates was transformed into E. coli and the recombinant E. coli
 isolates and used to transform other hosts e.g. prokaryotic cells, insect
 viruses, eukaryotic cells or plants, which could be used to control mites
 or, in the case of transgenic plants, confer resistance to mites.
 Query Match 1.33; Score 14; DB 12; Length 4155;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 402 accttactacggcg 415
 OY 355 ACCTTTACTACGCC 368
 ID 489
 RESULT
 ID Q0334 standard: DNA: 4155 BP.
 DT 26-WAN-1992 (first entry)
 B thuringiensis toxin gene ps17a.

were ampicillin; fluksicid; anthelmintic; parasite; as.
Bacillus thuringiensis.

DB 27-DEC-1991. 305047
PP 04-JUN-1980: US-572610
PP 24-JUL-1980: US-572746
PR 10-AUG-1980: US-556738
PR 17-JUL-1980: US-565344
PR 10-AUG-1980: US-565344
PR 27-MAR-1981: US-573772
PR 03-MAY-1981: US-593018
PA (MTCO.) MYCOGEN CORP.
PI Sarva KS, Payne JH, Schwab GE, Hickie LA, Galesan T;
PI MPEI 92-00106/01.
DR P-P&DN: R20066.
DE New Bacillus thuringiensis strains expressing coxins - have
PT been used to control nematodes, Helminths and flukes
PP e.g. liver fluke Fasciola hepatica
CC Claim 21, Page 24, 47pp: English.
CC The Psf1a toxin gene was isolated from B.t. strain PS17. Probe
CC QD019 was used to screen a PS17 total cellular DNA library. The
CC transformants into the acrylamide sensitive B.t. strain, HD-1 crys. The
CC recombinant B.t. strain was grown to sporulation and the toxin
CC protein purified by NaBr gradient centrifugation.
CC Genbank Accession # D01913.
SQ Sequence 4155 BP. 1386 A; 677 C; 811 G; 1381 T;
Query Match 1.38; Score 14; DB 3; Length 4155;
Best-Local Similarity 100.0%; Pres.No. 1;A+G+C; Indels 0; Gaps 0;
Ident 14; Conservative G; Mismatches 0;

DB 402 acccttaccgcggy 415
QT | | | | | | | | | |
| | | | | | | | | |
QT 355 ACCCTTACCACGC 368

RESULT 450
Q1400 standard; DNA, 4155 BP.
AC Q14000
DT 20-APR-1993 (first entry)
DE Bacillus thuringiensis gene 17a.
RW Toxin protein; ant. ds. strain PS17,
PF Bacillus thuringiensis
FF Location/Qualifiers
cds 1..4155
/*tag= a
/note= Protein 17a*

W09220802.A

[illegible]

```

CC This sequence represents the coding sequence for the splice variant of
CC the Drosophila melanogaster "timeless" (TIM) protein. The NTP is a nuclear
CC translocation protein (NTP) of the invention. The NTPs of the invention
CC are involved in circadian rhythms and cyclic transcription patterns related to sleep-wake
CC rhythms. The NTP is also preferably light sensitive, and has a stabilizing
CC effect on the circadian rhythm protein. The NTP also has the ability to
CC induce the NTPs of circadian rhythm entrainment to environmental cycles
CC of light. TIM is a nuclear protein which has homology to the family of
CC PER (PER) gene. PER is a nuclear protein which has homology to the family of
CC transcription factors containing the PAS domain, and is a protein
CC involved in circadian rhythms, but the biological function of PER is the
CC same as that of TIM. The NTPs of the invention are also phosphorylated with a circadian rhythm. The NTPs of the
CC invention, their fragments, agonists, and antagonists are useful
CC for preventing and/or treating disorders of a circadian rhythm such as
CC jet lag, shift work, and the like. The NTPs of the invention are also
CC detecting and/or measuring conditions so as to classify groups of
CC individuals with the disorders.
CC Sequence 4170 BP; 1078 A; 1128 C; 1102 G; 861 T;
CC
CC Query Match 1 33% Score 14; DB 39; Length 4170;
CC Best Local Similarity 100.0%; Pred. No. 3,474+02;
CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
DB 897 ttgacagagctgctg 905
Oy |||||
CC 84 ttccacgagctcac 97
CC
CC RESULT 492
CC ID G52123 standard; cDNA; 4175 BP.
CC AC G52123
CC DT 21-Nov-1994. (first entry)
CC OS Drosophila melanogaster
CC GN time-specific receptor tyrosine kinase; tek gene; mouse.
CC MU mouse; chromosome 4; cardiogenesis; as.
CC MS Mus musculus.
CC CD Location/Qualifiers
CC ccd 124..480
CC ccd /tag= a
CC ccd /product= tyrosine_kinase
CC
CC G22829391.A.
CC 14-DEC-1992; 0852931.
CC 30-JUL-1992; US-921795.
CC PA (MOON ) MOONSHI SHINJI HOSPITAL CORP.
CC FT FT 941-20254/element 2U, Hosanai J, Yamaguchi TP.

```


PT sequences
 PS Claim 1: Fig 2; 45pp; Chinese.
 CC Q5540 encodes R78519, the novel effector enzyme cardiac adenylyl
 CC kinase.
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1809 caagagacagta 1822
 Cp 86 CAGGACAGCACTA 73
 RESULT 498
 ID Q5540 standard; cDNA to mRNA; 4356 BP.
 DT 30-NOV-1995 (first entry)
 DE Brush-1 cDNA.
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;
 OS Homo sapiens.
 PH M09515334.A.
 PF 08-JUN-1995.
 PF 30-NOV-1994; U13823.
 PF 30-NOV-1994; U13823.
 PF 27-SEP-1994; U5-114598.
 PA (CALP.) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
 PI Chang J, Schott DR, Smith HS;
 PI WP1: 95-21528/28. from the Brush-1 tumour suppressor gene - useful
 PT as probes for detecting tumours and premalignant cells, esp. of the
 PT breast, and therapeutically
 PS Claim 2: Page 25-29; 39pp; English.
 CC cDNA obt'd. from breast tumor cell RNA was subjected to PCR using
 CC primers that flank the brush-1 gene. The resulting cDNA was
 CC sequenced. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence that can be used to detect this marker in breast tissue.
 CC Brush-1 nucleic acid may also be used for gene therapy in a
 CC subject deficient in a functional Brush-1 gene.
 CC Sequence 4356 BP; 1179 A; 352 G; 1373 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4062 aagttttttgag 4075
 Cp 1042 AAATGTTTCTGAG 1029

PT sequences
 PS Claim 1: Fig 2; 45pp; Chinese.
 CC Q5540 encodes R78519, the novel effector enzyme cardiac adenylyl
 CC kinase.
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1809 caagagacagta 1822
 Cp 86 CAGGACAGCACTA 73
 RESULT 498
 ID Q5540 standard; cDNA to mRNA; 4356 BP.
 DT 30-NOV-1995 (first entry)
 DE Brush-1 cDNA.
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;
 OS Homo sapiens.
 PH M09515334.A.
 PF 08-JUN-1995.
 PF 30-NOV-1994; U13823.
 PF 30-NOV-1994; U13823.
 PF 27-SEP-1994; U5-114598.
 PA (CALP.) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
 PI Chang J, Schott DR, Smith HS;
 PI WP1: 95-21528/28. from the Brush-1 tumour suppressor gene - useful
 PT as probes for detecting tumours and premalignant cells, esp. of the
 PT breast, and therapeutically
 PS Claim 2: Page 25-29; 39pp; English.
 CC cDNA obt'd. from breast tumor cell RNA was subjected to PCR using
 CC primers that flank the brush-1 gene. The resulting cDNA was
 CC sequenced. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence that can be used to detect this marker in breast tissue.
 CC Brush-1 nucleic acid may also be used for gene therapy in a
 CC subject deficient in a functional Brush-1 gene.
 CC Sequence 4356 BP; 1179 A; 352 G; 1373 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4062 aagttttttgag 4075
 Cp 1042 AAATGTTTCTGAG 1029

PT sequences
 PS Claim 1: Fig 2; 45pp; Chinese.
 CC Q5540 encodes R78519, the novel effector enzyme cardiac adenylyl
 CC kinase.
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1809 caagagacagta 1822
 Cp 86 CAGGACAGCACTA 73
 RESULT 498
 ID Q5540 standard; cDNA to mRNA; 4356 BP.
 DT 30-NOV-1995 (first entry)
 DE Brush-1 cDNA.
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;
 OS Homo sapiens.
 PH M09515334.A.
 PF 08-JUN-1995.
 PF 30-NOV-1994; U13823.
 PF 30-NOV-1994; U13823.
 PF 27-SEP-1994; U5-114598.
 PA (CALP.) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
 PI Chang J, Schott DR, Smith HS;
 PI WP1: 95-21528/28. from the Brush-1 tumour suppressor gene - useful
 PT as probes for detecting tumours and premalignant cells, esp. of the
 PT breast, and therapeutically
 PS Claim 2: Page 25-29; 39pp; English.
 CC cDNA obt'd. from breast tumor cell RNA was subjected to PCR using
 CC primers that flank the brush-1 gene. The resulting cDNA was
 CC sequenced. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence that can be used to detect this marker in breast tissue.
 CC Brush-1 nucleic acid may also be used for gene therapy in a
 CC subject deficient in a functional Brush-1 gene.
 CC Sequence 4356 BP; 1179 A; 352 G; 1373 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4062 aagttttttgag 4075
 Cp 1042 AAATGTTTCTGAG 1029

PT sequences
 PS Claim 1: Fig 2; 45pp; Chinese.
 CC Q5540 encodes R78519, the novel effector enzyme cardiac adenylyl
 CC kinase.
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1809 caagagacagta 1822
 Cp 86 CAGGACAGCACTA 73
 RESULT 498
 ID Q5540 standard; cDNA to mRNA; 4356 BP.
 DT 30-NOV-1995 (first entry)
 DE Brush-1 cDNA.
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;
 OS Homo sapiens.
 PH M09515334.A.
 PF 08-JUN-1995.
 PF 30-NOV-1994; U13823.
 PF 30-NOV-1994; U13823.
 PF 27-SEP-1994; U5-114598.
 PA (CALP.) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
 PI Chang J, Schott DR, Smith HS;
 PI WP1: 95-21528/28. from the Brush-1 tumour suppressor gene - useful
 PT as probes for detecting tumours and premalignant cells, esp. of the
 PT breast, and therapeutically
 PS Claim 2: Page 25-29; 39pp; English.
 CC cDNA obt'd. from breast tumor cell RNA was subjected to PCR using
 CC primers that flank the brush-1 gene. The resulting cDNA was
 CC sequenced. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence that can be used to detect this marker in breast tissue.
 CC Brush-1 nucleic acid may also be used for gene therapy in a
 CC subject deficient in a functional Brush-1 gene.
 CC Sequence 4356 BP; 1179 A; 352 G; 1373 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4062 aagttttttgag 4075
 Cp 1042 AAATGTTTCTGAG 1029

PT sequences
 PS Claim 1: Fig 2; 45pp; Chinese.
 CC Q5540 encodes R78519, the novel effector enzyme cardiac adenylyl
 CC kinase.
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1809 caagagacagta 1822
 Cp 86 CAGGACAGCACTA 73
 RESULT 498
 ID Q5540 standard; cDNA to mRNA; 4356 BP.
 DT 30-NOV-1995 (first entry)
 DE Brush-1 cDNA.
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;
 OS Homo sapiens.
 PH M09515334.A.
 PF 08-JUN-1995.
 PF 30-NOV-1994; U13823.
 PF 30-NOV-1994; U13823.
 PF 27-SEP-1994; U5-114598.
 PA (CALP.) CALIFORNIA PACIFIC MEDICAL CENT RES INST.
 PI Chang J, Schott DR, Smith HS;
 PI WP1: 95-21528/28. from the Brush-1 tumour suppressor gene - useful
 PT as probes for detecting tumours and premalignant cells, esp. of the
 PT breast, and therapeutically
 PS Claim 2: Page 25-29; 39pp; English.
 CC cDNA obt'd. from breast tumor cell RNA was subjected to PCR using
 CC primers that flank the brush-1 gene. The resulting cDNA was
 CC sequenced. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence. Isolated clones were used to assemble a Brush-1 cDNA
 CC sequence that can be used to detect this marker in breast tissue.
 CC Brush-1 nucleic acid may also be used for gene therapy in a
 CC subject deficient in a functional Brush-1 gene.
 CC Sequence 4356 BP; 1179 A; 352 G; 1373 T;
 CC Gaps 0;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4062 aagttttttgag 4075
 Cp 1042 AAATGTTTCTGAG 1029

RESULT 501
ID T05556 standard; DNA: 4405 BP.
AC T05556;
AT 14-APR-1996 (first entry)
DT 14-APR-1996 (first entry)
KW human PACE-encoding sequence.
KW subtilisin-like protease; precursor polypeptide processing;
KW blood coagulation protein precursor; da.
OS Homo sapiens.
RN 5'utr Location/Qualifiers
FT 88..107
FT signal_peptide 408..485
FT cda /*tag= a
FT /*tag= b
FT /*tag= c
FT mat_peptide 486..2789
FT /*tag= d PACE
FT 3'utr 2'product=1
FT poly_a_signal 4345..4350
FT /*tag= e
FT /*tag= f
FT US640950-A
PN 20-NOV-1992; 885972.
PR 20-NOV-1992; US-621092.
PR 29-NOV-1990; US-621443.
PR 29-NOV-1990; US-621443.
PR 10-NOV-1990; US-621457.
PR 10-NOV-1990; US-621457.
PR 20-MAY-1992; US-885972.
PA (CHIR) CHIRON CORP.
PA (GEMT) GENETICS INST. INC.
PI Wong PA; Blake AJ; Kaufman RJ; Tekamp-Olsen P; Wasley L;
WP1: 95-373215/48.
DR P-PSDB: R77540.
DT Host cells expressing PACE for processing/conversion of precursor
FT Host cells expressing PACE for processing/conversion of precursor
FT biological activity and blood coagulation proteins.
PS Claim 7; Fig 2A-2F; 50pp; English.
CC The DNA encodes PACE, a paired basic amino acid converting enzyme,
CC which has structural homology with a subtilisin-like serine
CC protease. The DNA is expressed in a host cell (yeast, mammalian or insect preferably
CC CHO cells) along with a heterologous polynucleotide encoding a
CC precursor polypeptide (preferably a precursor of a blood coagulation
CC factor or a protein requiring gamma-carboxylation for biological
CC activity). The precursor polypeptide is converted to the mature
CC form of basic amino acids (Lysine, Lysine, Arginine) by the co-expressed

CC PACE into its mature form. Thus, the host cell is capable of
CC producing high levels of PACE and active, mature heterologous
CC proteins. The active, mature heterologous proteins are PACE,
CC protein-C, protein-S, Factor-X, Factor-VII and bone-gamma
CC carboxylglutamate protein.
SQ See T05555.
SQ Sequence 4405 BP; 823 A; 1442 C; 1329 G; 811 T;
Query Match 1.3%; Score 14; DB 17; Length 4405;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3437 ctgaccctctgct 3450
QY 904 CTGACCTCTGCT 917
|||||
RESULT 502
ID Q46293 standard; DNA: 4415 BP.
AC Q46293;
AT 09-DEC-1993 (first entry)
DT 09-DEC-1993 (first entry)
KW Regulation; Phenylalanine ammonia lyase; PAL; pal; as.
OS Pisum sativum.
PN J05153978-A.
PR 22-NOV-1991; 317126.
PR 29-NOV-1991; JP-317126.
PA (TAKS) TAKASAGO PERFUMERY CO LTD.
PR WP1: 93-231495/29.
PT Regulator gene of specified sequence - and phenylalanine ammonia
PT Regulator gene of specified sequence - and phenylalanine ammonia
PT Claim 3; Page 7-8; 9pp; Japanese.
CC The regulatory fragment of the phenyl alanine ammonia lyase gene is
CC capable of artificially accelerating or suppressing the transcription of
CC the structural gene of a host plant.
SQ Sequence 4415 BP; 1415 A; 725 C; 793 G; 1482 T;
Query Match 1.3%; Score 15; DB 7; Length 4415;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3705 caagtgtctaaagg 3719
QY 327 CAACTGCTAAAGG 341
|||||
RESULT 503
ID T03080 standard; DNA: 4425 BP.
AC T03080;
AT 14-FEB-1996 (first entry)
DT 14-FEB-1996 (first entry)

DE Protein tyrosine-kinase SAL-S1 gene.
KW Protein tyrosine-kinase; PK; SAL-S1; agonist; cell growth;
KW differentiation; ss.
RN 5'utr Location/Qualifiers
FT 30..3927
FT signal_peptide 30..102
FT mat_peptide 103..1924
FT /*tag= c
FT W0927061-A1.
PN 12-DEC-1995; U04228.
PR 04-APR-1994; US-222616.
PA (GETH) GENETECH INC.
PI Bennett BD; Goeddel D; Lee JM; Matthews W; Tsai SP;
DR WP1: 95-365140/47.
DR P-PSDB: R85937.
DT Agonist antibodies which activate specific protein tyrosine
FT kinases(s) also activate chimeric proteins of kinase extracellular
FT domain of cytochrome c and a kinase domain. The kinase domain is
FT mutation of cytochrome c domain. The kinase domain is used for studying, and therapeutic
PS Disclosure; Fig 15A-F; 125pp; English.
CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used
CC to screen cDNA libraries to identify novel PTK genes. The SAL-S1
CC cDNA library was screened for several regulatory cell
CC fragments. The gene can be used to produce recombinant SAL-S1
CC fragments, to detect related genes, and to design drugs. Peptides
CC or antisense nucleotides that modulate PTK activity.
SQ Sequence 4425 BP; 939 A; 1348 C; 1361 G; 777 T;
Query Match 1.3%; Score 14; DB 16; Length 4425;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 4051 ccagagagagagag 4064
Cp 160 CCAGAGAGAGAG 147
|||||
RESULT 504
ID T89487 standard; cDNA: 4440 BP.
AC T89487;
DT 19-FEB-1998 (first entry)
DT 19-FEB-1998 (first entry)
KW Human A20 protein cDNA.
KW NF-kappa-B inhibitor; organ transplantation; ischaemia;
KW inflammation; gene therapy; endothelial cell; da.
OS Homo sapiens.

FN Key Location/Qualifiers
FT CDS 67..2439
FT /*tag= a
FT W0370083-A1.
PN 21-APR-1997; 200676.
PR 13-FEB-1997; 200676.
PR 19-APR-1996; US-634995.
PR 14-FEB-1996; US-601515.
PA (BOWE) BOWEN HOSPITAL DEACONES HOSPITAL.
PI Bach FB; Perran C;
PR WP1: 97-424975/39.
DR P-PSDB: W31528.
DT Recombinant endothelial cell containing DNA encoding anti-apoptotic
FT protein tyrosine-kinase (PTK) domain. The PTK domain is
FT useful for generating tissues or organs for transplantation
PS Example 8; Page 40-44; 75pp; English.
CC This cDNA clone encodes human A20 (see W31528), a protein capable
CC of blocking or suppressing NF-kappa B (NF-kB) activation. It was
CC obtained by cloning a cDNA library from a human endothelial cell
CC susceptible to an inflammatory or other immunological stimulus
CC comprising inserting into the cell, DNA encoding an anti-apoptotic
CC protein tyrosine-kinase (PTK) domain. The PTK domain is used for
CC the cellular activating stimulus. Suitable anti-apoptotic proteins
CC include A20, BCL-2 (see W31528), BCL-XL (see W31530) and A1 (see
CC W31531) and their deletion mutants. Also claimed are: (1) a
CC non-human transgenic animal, (2) a method of generating a transgenic
CC animal, (3) a method of generating a transgenic animal, (4) a
CC method of generating a transgenic animal, (5) a method of generating a
CC method of generating a transgenic animal, (6) a method of generating a
CC tissues or organs for transplantation into recipient species.
SQ Sequence 4440 BP; 1206 A; 1070 C; 1055 G; 1109 T;
Query Match 1.3%; Score 15; DB 35; Length 4440;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2483 scgtcgtcgtcgtcgt 2497
QY 492 CCGTCGTCGTCGTCAT 506
|||||
RESULT 505
ID T64783 standard; cDNA: 4503 BP.
AC T64783;
DT 03-SEP-1997 (first entry)
DT 03-SEP-1997 (first entry)
KW Oxygen regulated protein ORP 150 cDNA.
KW hypoglycaemia; hypoxia; gene therapy; diagnosis; da.

US-08-887-977-9.mg

```

FT      /tag= c
FT      1728..2091
intron
FT      /tag= d
FT      2092..2271
exon
FT      /tag= e
FT      2272..2375
intron
FT      /tag= f
FT      2376..2673
exon
FT      /tag= g
FT      2674..2856
intron
FT      /tag= h
FT      2857..2907
exon
FT      /tag= i
FT      2908..2994
intron
FT      /tag= j
FT      2995..3251
exon
FT      /tag= k
FT      3252..3340
intron
FT      /tag= l
FT      3341..3575
exon
FT      /tag= m
FT      3576..3670
intron
FT      /tag= n
FT      3671..3792
exon
FT      /tag= o
J01242687.A.
17-SEP-1989.
14-MAR-1989; JP-061702.
(MOCH) MOCHIDA PHARM KK.
WFL: 50-364257/49.
Plasmid and expression plasmid - contains promoter region of
human polypeptide chain elongation factor
Disclosure: Fig 3; 23pp; Japanese.
This patent document discloses a DNA sequence may be used as an expression
vector or plasmid in transgenic animals. This plasmid has been modified
over plasmids using the SV40 promoter, stable for a month in cells,
and able to express a wide range of products.
Sequence 4691 BP; 1200 A; 990 C; 1233 G; 1268 T;
Score 14; DB 1; Length 4691;
Query Local Similarity 100.0%;
Matches 14; Conservative O; Mismatches O; Indels O; Gaps:
DB      2991 gptttcttattccc 2304
          |||||
QY      637 GGTTTCCTTATCCC 650

```

Page 505

22-MAR-1996; EU1271.
 PP (SANO) SARDZ 050597.
 PA (SANO) SARDZ PATENT QMBH.
 PA (SANO) SARDZ-ERFINDUNGEN VERW QMBH.
 PA WPI: 96-443194/44.
 PP Reombinant nucleic acid contg. modified tospovirus sequence - used to prepare plants resistant or tolerant to tospovirus
 CC A DNA 1; Page 173810; 38pp; English.
 CC A DNA 1; Page 173810; 38pp; English.
 CC The tomato spotted wilt virus M-RNA sequence to one which encodes the putative viral movement protein. A recombinant nucleic acid contains a transcriptional regulatory region and, under its control, a coding sequence for the movement protein. The expression or inactivation of the start codon, and insertion or deletion of nucleotides 3' of the start codon to alter the reading frame of the encoded RNA. A plant having such a construct stably integrated and expressed in its genome is resistant to, or tolerant of, tospovirus.
 CC Tospovirus.
 CC Sequence 4821 BP; 1569 A; 850 C; 870 G; 1532 T; Query Match
 CC Seq. Similarity 100.0%; 100.0%; 100.0%; 100.0%;
 CC Matches 14; Conservative
 DB 2914 tgcctttgaatt 2927
 CP 722 tgcctttgaatt 709

SEQUENCE 512
 ID: T05848 standard; DNA: 4951 BP.
 AC T05848;
 CD 03-APR-1996 (first entry)
 DE Polyelectronic cps gene locus operon of S.pneumoniae
 DE Polyelectronic cps gene locus operon of S.pneumoniae
 DE serotype; diagnosis; prevention; Streptococcus pneumoniae; detection;
 DE Streptococcus pneumoniae.
 FE Key location/Qualifiers
 FE cds /tag= a
 FE /product= Capsular polysaccharide CpsD.
 FT cds 1377..2543
 FT /tag= b
 FT /product= Capsular polysaccharide CpsE.
 FT cds 2707..3761
 FT /tag= c
 FT /product= Capsular polysaccharide CpsF.
 FT cds 4821..5811
 FT /tag= d
 FT /product= Capsular polysaccharide CpsG.

FT W0951246-A1.
 PN 13-OCT-1994.
 PP 05-APR-1994: D03708.
 PR 16-MAY-1994: US-243546.
 PA (UABR-) UAB RES FOUND.
 PI Dillard JP, Rother J.
 DI 01-APR-1997: D06087.
 DR P-PSDB: R81039, R83040, R83041.
 DT New streptococcus pneumoniae capsular polysaccharide genes - used
 PT for detection, serotyping and for diagnosis and prevention of S.
 CC pneumoniae infection.
 CC Sequences encoding the 5' flanking region of the capsular
 CC polysaccharide gene (cps) of streptococcus pneumoniae and which are
 CC of sufficient length to allow hybridisation under standard
 CC hybridisation conditions to a S. pneumoniae cps gene flanking region
 CC nucleotide sequence. The sequences are useful for diagnosis and
 CC may also be used for the diagnosis and prevention of S. pneumoniae
 CC infection.
 SQ Sequence 4951 BP: 1565 A; 749 C; 1028 G; 1609 T;
 Query Match 1.38; Score 14; DB 17; Length 4951;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3032 tgaagattattttg 3045
 QY 17 TCAGATTATTG 30

RESULT 513
 ID Q73445 standard; cDNA: 4975 BP.
 AC Q73445;
 DT 09-MAY-1995 (first entry)
 DE IGF-1 receptor.
 KW IGF-1 receptor; insulin-like growth factor-I receptor;
 KW cell proliferation; Hela; mutagenesis; tyrosine kinase;
 KW phosphorylation; probe; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 32..4135
 FT /tag= a
 FT signal_peptide 32..121
 FT /tag= b
 FT mat_peptide 121..4135
 FT /tag= c
 FT /product= IGF-1 receptor alpha subunit
 FT mat_peptide 2552..4132
 FT /tag= d
 FT /product= IGF-1 receptor beta subunit

DR WPI: 96-259475/26.
 DR P-PSDB: R95244.
 DT Induction of resistance to tumour cell growth - by insertion of
 PT diffusion chamber contg. tumour cell culture supplemented with, e.g.
 CC a diffusible growth factor.
 CC Discloure: Fig 4A-G; 47pp; English.
 CC The nucleotide sequence (729608) coding for insulin growth factor
 CC 1 receptor (IGF1R, R95244) was used for the design of antisense
 CC oligonucleotides (IGF1R-1 and IGF1R-2) complementary to codons 1-308 of
 CC the mature protein.
 CC respectively. The antisense sequences act as apoptosis-inducing
 CC agents when supplied to tumour cells growing in a diffusion chamber.
 CC Induction of resistance to tumour growth occurs upon insertion
 CC of the chamber into a mouse.
 SQ Sequence 4989 BP: 1216 A; 1369 C; 1322 G; 1082 T;
 Query Match 1.38; Score 14; DB 20; Length 4989;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 188 gctctgagaactgc 201
 CP 111 CCTCTGAGACTGC 98

RESULT 515
 ID T13329 standard; cDNA to mRNA: 4989 BP.
 AC T13329;
 DT 03-SEP-1996 (first entry)
 DE Human type I insulin-like growth factor receptor cDNA.
 KW Insulin-like growth factor 1 receptor; IGF 1R; somatomedin C;
 KW smooth muscle; cell growth; cell proliferation; healing;
 KW tumour regression; angiogenesis; antisense RNA; atherosclerosis;
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 46..4149
 FT /tag= a
 FT signal_peptide 46..135
 FT /tag= b
 FT mat_peptide 136..4146
 FT /tag= c
 FT /product= IGF-1 receptor
 PN W09610401-A1.
 PP 11-APR-1996.
 PR 27-SEP-1995: U12563.
 PR 04-OCT-1994: US-317898.
 PA (LDBU-) LDBUNGEN INC.
 PI Baserga RL, O'Connor R.
 DI 01-APR-1997: D06087.
 DR P-PSDB: R91429.
 DT Insulin-like growth factor 1 receptor antisense RNA and

PN W09420304-A.
 PN 13-OCT-1994.
 PP 05-APR-1994: D03708.
 PR 16-MAY-1994: US-243546.
 PA (UABR-) UAB RES FOUND.
 PI Dillard JP, Rother J.
 DI 01-APR-1997: D06087.
 DR P-PSDB: R81039, R83040, R83041.
 DT New streptococcus pneumoniae capsular polysaccharide genes - used
 PT for detection, serotyping and for diagnosis and prevention of S.
 CC pneumoniae infection.
 CC Sequences encoding the 5' flanking region of the capsular
 CC polysaccharide gene (cps) of streptococcus pneumoniae and which are
 CC of sufficient length to allow hybridisation under standard
 CC hybridisation conditions to a S. pneumoniae cps gene flanking region
 CC nucleotide sequence. The sequences are useful for diagnosis and
 CC may also be used for the diagnosis and prevention of S. pneumoniae
 CC infection.
 SQ Sequence 4951 BP: 1565 A; 749 C; 1028 G; 1609 T;
 Query Match 1.38; Score 14; DB 17; Length 4951;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 174 gctctgagaactgc 187
 CP 111 CCTCTGAGACTGC 98

RESULT 514
 ID T29608 standard; DNA: 4989 BP.
 AC T29608;
 DT 12-AUG-1996 (first entry)
 DE IGF-1 receptor gene.
 KW Antisense RNA; tumour; cancer; apoptosis; diffusion chamber;
 KW insulin growth factor 1 receptor; IGF1R; as.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 36..4149
 FT /tag= a
 FT signal_peptide 36..135
 FT /tag= b
 FT mat_peptide 136..4146
 FT /tag= c
 FT /product= IGF-1 receptor
 PN W09614746-A1.
 PP 15-NOV-1995.
 PR 15-NOV-1995: U14952.
 PR 16-NOV-1994: US-340732.
 PA (UTZE-) UNIV JEFFERSON THOMAS.
 PI Abraham D, Baserga R, Resnicoff M;

PT ATG-directed sense oligonucleotide(s) - useful for regulating
 PT growth factor receptor gene expression for e.g. wound healing and
 PT atherosclerosis.
 PS Discloure: Page 46-52; 73pp; English.
 CC A cDNA clone (T13329) codes for human insulin-like growth factor
 CC receptor type I (IGF1R). The cDNA was used to generate a rat
 CC vascular smooth muscle cell (VSMC) proliferative response. Rat
 CC IGF 1R cDNA (T13330) has provided the basis for antisense methods
 CC (see T13325 and T13328) to down-regulate IGF 1R gene expression and
 CC retard VSMC growth, e.g. to treat restenosis and atherosclerosis,
 CC and to stimulate VSMC growth e.g. for healing wounds and burns.
 CC Methods to stimulate VSMC growth e.g. for healing wounds and burns.
 SQ Sequence 4989 BP: 1216 A; 1369 C; 1322 G; 1082 T;
 Query Match 1.38; Score 14; DB 19; Length 4989;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 188 gctctgagaactgc 201
 CP 111 CCTCTGAGACTGC 98

RESULT 516
 ID T13329 standard; cDNA: 4989 BP.
 AC T13329;
 DT 14-APR-1998 (first entry)
 DE Human insulin-like growth factor 1 receptor cDNA.
 KW Insulin-like growth factor 1 receptor; IGF-1R; apoptosis;
 KW smooth muscle; cell growth; cell proliferation; healing;
 KW tumour regression; angiogenesis; antisense RNA; atherosclerosis;
 KW therapy; human; as.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 46..4149
 FT /tag= a
 FT signal_peptide 46..135
 FT /tag= b
 FT mat_peptide 136..4146
 FT /tag= c
 FT /product= IGF-1 receptor
 PN W09737010-A1.
 PP 09-OCT-1997.
 PR 01-APR-1997: D06087.
 PR 01-APR-1997: D06087.
 PA (LDBU-) LDBUNGEN INC.
 PI Baserga RL, O'Connor R.
 DI 01-APR-1997: D06087.
 DR P-PSDB: R91429.
 DT Insulin-like growth factor 1 receptor -
 PT especially C-terminally truncated or derived peptide(s), useful for

FT		/note= "g polymorphic HaeIII site"
FT		5035..3904
FT	misc_feature	/tag= h
FT		5035..3904
FT		/note= "g polymorphic MspI site"
FT		5035..3904
FT	misc_feature	/tag= i
FT		5035..3904
FT		/note= "non-depleted 3.5 kb intron sequence"
FT	misc_feature	4012..4013
FT		4012..4013
FT		/note= "non-depleted 0.4 kb intron sequence"
FT	misc_feature	4567..4568
FT		4567..4568
FT		/tag= k
FT		1025..1439
FT	misc_signal	/tag= l
FT		1025..1439
FT		/note= "13 nucleotide element"
FT	misc_signal	1489..1501
FT		1489..1501
FT		/note= "13 nucleotide element"
FT	TATA_signal	1546..1551
FT		1546..1551
FT		/tag= n
FT		5025..5031
FT	polya_signal	/note= "putative TATA-box"
FT		5025..5031
FT		/tag= o
FT		/note= "putative polyadenylation signal"
FT		50373589-52.
FT	ORF73589-52.	
FT		27-MAR-1997; NM157.
FT		28-MAR-1996; EP-200855.
PR	(DALL)	DALLAND BV.
PR	(DALL)	DALLAND BV.
PA	(ENVE)	NORD NEDERLANDS EN DIERGEZONDHEID ID-D.
PA	(ENVE)	NORD NEDERLANDS VARENSSTAMBOEK BV.
PA	(PROV.)	PROVA BV.
PA	(STAN.)	STAMBOEK ZUID BV.
PA		WIPB097-469563/A5.
DR	P-FSDM:	W1534.
DR		WIPB097-469563/A5.
PT	Pig heart fatty acid-binding protein gene - used to identify	
PT	polymorphisms associated with production traits, e.g. body weight,	
PT	polymorphisms associated with production traits, e.g. body weight,	
PS	Gla1n 1: Fig 1: 41pp; Enghash.	
PS	The present sequence encodes porcine heart-fatty acid binding	
CC	protein (H-FABP). The H-FABP gene can be used to localise, identify	
CC	essential parts of the genome which are associated with production traits,	
CC	especially where these are associated with production traits.	
CC	Alleles of the porcine H-FABP gene can be marked, allowing them to	
CC	be distinguished, preferably by detection of specific restriction	
CC	sites, e.g. MspI, HaeIII or KpnI. The H-FABP genes can also be used	
CC	to study the inheritance of production traits in pig breeds. By	
CC	loci in samples by amplification of specific genomic fragments. By	

22-OCT-1982; 05-955055
PA (TEMA) UNIV TEXAS SYSTEM
PR Baseman JG, Ballo SF, Su CJ;
PI WP1: 95-041662/02.
PT Detecting M. pneumoniae contamination or infection - using DNA
PT sequencing.
PT Detecting M. pneumoniae pl polypeptide as a hybridisation probe to
PT detect pathogenic Mycoplasmas
PT Disclosure; fig. 6A-6F; 47pp; English.
PT Disclosed by: M. J. Baseman, University of Texas at Dallas, Texas
PT base-harboring Mycoplasma pneumoniae M179 oligonucleotide probes
CC based on the 5'-terminal sequence of pl were used in Southern blots
CC to isolate pl-encoding DNA, which was then cloned and sequenced
CC (JQ3074); the predicted amino acid sequence is given in R87518.
CC Sequence 5168 BP; 1384 A; 1455 C; 1217 G; 1112 T;
CC

Query Match 1.3%; Score 14; DB 13; Length 5168;
Best Local Similarity 100.0%; Freq. No. 3,476/02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1675 gaatggcagatggt 1698
| | | | | | | | | | | | | | | |
Cp 245 GCAATGCCAGTGT 232

RESULT 521
ID T0801 standard; DNA: 5169 BP.
T0001
D 06-JUN-1996 (first entry)
DE Streptomyces vedemorensis fosfomycin biosynthesis related genes.
DE Fosfomycin biosynthesis; methylation; epoxidation; enzyme;
DE Streptomyces lividans; ds. recombinant production;
KW Streptomyces lividans; ds.
OS Streptomyces vedemorensis.
FH Location/Qualifiers
FT /size: 4216
FT /product: methylation enzyme (R85167)
FT /tag: b
FT 4233..4829
cda /product: epoxidation enzyme (R85168)
J07265080-A.
17-OCT-1995.
PA 29-MAR-1994; 059844
PR 29-MAR-1994; 059844
PI (MELJ) MELJ1; GENEX MAISHA LTD.
PT WP1: 95-388688/50.
PT P-DSB3; R85167; R85168.
DR Fosfomycin biosynthesis related genes - encoding methylation and
DR epoxidation; recombinant production of fosfomycin.
Citation 2: Pages 7-11; 13pp; Japanese.
P5

PR 05-MAR-1991; US-665792.
 PA (TEMA) UNIV TEXAS.
 DE M. pneumoniae strain 1688, SU CJ;
 WI: 94-04285/05.
 DR P-PSDB: R47811.
 FT New antigenic fragments of Mycoplasma pneumoniae cytoadhesin -
 useful in protective vaccines and as diagnostic immunoassay
 reagents.
 Example 2; Fig 6; 52pp; English.
 CC DNA was extracted from M. pneumoniae M139, fragmented with
 endonucleases and a 3kb HindIII fragment was identified which
 K-terminally with oligonucleotides correspond to the known
 K-terminus of the cytoadhesin protein. The fragment was
 clone 62A. An EcoRI-PstI fragment from this clone was used to
 screen an EcoRI digest of M. pneumoniae DNA to identify a 6kb
 fragment containing the whole gene. The insert was subcloned and
 sequenced (527). The cytoadhesin protein encoded by the
 insert is identical for all but one residue to the known
 sequence. 5169 BP; 1326 A; 1454 C; 1278 G; 1111 T;
 Query Match
 ID Q1252 standard; DNA; 5169 BP.
 AC Q12520.
 DE M. pneumoniae P1 cytoadhesin protein.
 KW Vaccine; diagnosis; cytoadhesin; cytoadherence; antibodies; ds.
 PS Mycoplasma pneumoniae strain M139.
 FT cds Location/Qualifiers
 FT 159..5039
 FT /tag= a
 FT /product= P1 protein
 FT /note= "pref. fragment"
 FT /tag= b
 FT /tag= c
 FT /tag= d
 FT /note= "pref. fragment"
 FT /tag= e
 FT /tag= f

CC The S. wedmorensis fosfomycin biosynthesis related gene T06001, and
 CC its promoter, respectively. By cloning and sequencing of the M1
 CC and EP enzyme genes in S. wedmorensis N57/PB623 (FERN P-13840),
 CC its plasmid pF623 can be used to prepare the S. lividans variant
 CC pF623 (FERN P-1243). The prep. variant can be used to prepare
 CC the fosfomycin variant. The fosfomycin variant can be used for
 CC the study of the role of the fosfomycin biosynthesis genes in
 CC the prevention of mycoplasma infection. The DNA can also be used for
 CC detection and isolation of other mycoplasma DNA, e.g. from M.
 Query Match
 ID Q5775 standard; DNA; 5169 BP.
 AC Q57750.
 DE Mycoplasma pneumoniae P1 cytoadhesin structural gene.
 KW anti-mycoplasma pneumoniae; P1 cytoadhesin; antigen; immunogen; epitope;
 PS Mycoplasma pneumoniae strain M139.
 FT cds Location/Qualifiers
 FT 122..127
 FT /tag= a
 FT /note= "not indicated on Fig. 6"
 FT /tag= b
 FT /tag= c
 FT /tag= d
 FT /product= P1 cytoadhesin
 FT /note= "except pos:4929..4931; aa:Pro
 FT /note= "in Mycoplasma sp., T06 codons encode
 FT 5043..5169
 FT /tag= e
 FT /tag= f
 FT /tag= g
 FT /tag= h
 FT /tag= i
 FT /tag= j
 FT /tag= k
 FT /tag= l
 FT /tag= m
 FT /tag= n
 FT /tag= o
 FT /tag= p
 FT /tag= q
 FT /tag= r
 FT /tag= s
 FT /tag= t
 FT /tag= u
 FT /tag= v
 FT /tag= w
 FT /tag= x
 FT /tag= y
 FT /tag= z
 FT /tag= A
 FT /tag= B
 FT /tag= C
 FT /tag= D
 FT /tag= E
 FT /tag= F
 FT /tag= G
 FT /tag= H
 FT /tag= I
 FT /tag= J
 FT /tag= K
 FT /tag= L
 FT /tag= M
 FT /tag= N
 FT /tag= O
 FT /tag= P
 FT /tag= Q
 FT /tag= R
 FT /tag= S
 FT /tag= T
 FT /tag= U
 FT /tag= V
 FT /tag= W
 FT /tag= X
 FT /tag= Y
 FT /tag= Z
 FT /tag= [
 FT /tag= \]
 FT /tag= ^
 FT /tag= _
 FT /tag= `

CC genitalium.
 CC See also Q12518-20.
 90 Sequence 5169 BP; 1323 A; 1454 C; 1279 G; 1113 T;
 Query Match
 ID Q12518 standard; DNA; 5169 BP.
 AC Q125180.
 DE M. pneumoniae P1 cytoadhesin protein.
 KW Vaccine; diagnosis; cytoadhesin; cytoadherence; antibodies; ds.
 PS Mycoplasma pneumoniae strain M139.
 FT cds Location/Qualifiers
 FT 159..5039
 FT /tag= a
 FT /product= P1 protein
 FT /note= "pref. fragment"
 FT /tag= b
 FT /tag= c
 FT /tag= d
 FT /note= "pref. fragment"
 FT /tag= e
 FT /tag= f

FT misc_rna /note= "pref. fragment"
 FT 4368..5039
 FT /tag= a
 FT /tag= b
 FT /tag= c
 FT /tag= d
 FT /tag= e
 FT /tag= f
 FT /tag= g
 FT /tag= h
 FT /tag= i
 FT /tag= j
 FT /tag= k
 FT /tag= l
 FT /tag= m
 FT /tag= n
 FT /tag= o
 FT /tag= p
 FT /tag= q
 FT /tag= r
 FT /tag= s
 FT /tag= t
 FT /tag= u
 FT /tag= v
 FT /tag= w
 FT /tag= x
 FT /tag= y
 FT /tag= z
 FT /tag= A
 FT /tag= B
 FT /tag= C
 FT /tag= D
 FT /tag= E
 FT /tag= F
 FT /tag= G
 FT /tag= H
 FT /tag= I
 FT /tag= J
 FT /tag= K
 FT /tag= L
 FT /tag= M
 FT /tag= N
 FT /tag= O
 FT /tag= P
 FT /tag= Q
 FT /tag= R
 FT /tag= S
 FT /tag= T
 FT /tag= U
 FT /tag= V
 FT /tag= W
 FT /tag= X
 FT /tag= Y
 FT /tag= Z
 FT /tag= [
 FT /tag= \]
 FT /tag= ^
 FT /tag= _
 FT /tag= `

```

Query Match      138; Scores 14; DP 2; Length 5186;
Best Local Similarity 100.0%; Pred No. 3,47+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1387 tctgttttcgctt 7400
QY 755 TCTGTTTCGCTTGGCT 1400

          RESUVR   525
ID    TG333 standard; DWA, 5198 bp.
AC    TG0333;
DE    28-MAY-1995 (first entry)
DT     TM gene
DD     TM gene
DS     TM gene
OS     TM gene
KW     sleep-wake cycle; light sensitive; circadian rhythm;
NM     Drosophila period gene; transcription factor; PAS domain; depression;
NR     neurolept analgesia;
NS     Drosophila melanogaster;
NC     DRUGS;
CC     RPT;
FT     exon                                243..3555/Qualifiers
                                           /start=1
                                           /end=3555
                                           /number=1
cds                                     243..3611
                                           /start=1
                                           /end=3611
                                           /number=1
misc_feature                          2387..2450
                                           /product=unspliced TIM
                                           /start=c
                                           /end=c
                                           /deleted=in tim O1 variant*
intron                               3556..3792
                                           /start=d
                                           /end=d
exon                                 3793..4646
                                           /number=1
                                           /start=e
                                           /end=e
                                           /number=2
WG9629A06-A2.
26-SEP-1996. PD
20-MAY-1996. UD8305.
16-MAY-1995. US-442214.
02-NOV-1995. US-552354.
(DIPE.) UNIV PENNSYLVANIA.
OTIO UNIV ROCHESTER.
HPI: 96.44152/74.
PDBD: W14052.
R nuclear translocation protein which binds to protein involved in
or treatment of disorders associated with circadian rhythms, e.g.
jetlag and narcolepsy.

```

```

Nov 17 08:55:26 1998

US-08-087-977-9.2mg

CC This repetitive element is used to synthesize DNA primers which may
CC be used for testing for infestation of plant tissue with L. maculana,
CC particularly with respect to blackleg infestation of Cruciferae,
CC especially rape (Brassica napus). Infestation testing may be
CC performed in the field in about 30 minutes per haploid
CC genome only in virulent isolates of L. maculana and is present on
CC every chromosome.
CC
CC Sequence 5238 bp: 1921 A; 836 C; 939 G; 1542 T;
CC
CC Query Match 133; Score 14; DB 13; Length 5238;
CC Best Local Similarity 100.0%; Pred. No. 3 474+02;
CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC bp 3426 gccacgagagacc 3439
CC cp 1019 gccacgagagacc 1006
CC
CC RESULT 527
CC ID DB QBL85 standard; DNM: 5248 BP.
CC AC QBL85;
CC DT 09-NOV-1995 (first entry)
CC DE Bacillus circulans CFase complete gene sequence.
CC EF Cyclusinulooligosaccharide fructanotransferase; Bacillus circulans;
CC FJ Cyclusinulooligosaccharide fructanotransferase; Bacillus circulans;
CC FM Cyclic inulooligosaccharide fructanotransferase; Insulin; ds.
CC GN Bacillus circulans MC1-2554.
CC PH Location/Qualifiers
CC PS cds
CC FT
CC FT /product= cyclic-inulooligosaccharide
CC FT fructanotransferase
CC FT misc_difference 4954
CC FT
CC FT /tag= b
CC FT reproduced in specification as P
CC misc_difference 4954, 4956
CC FT
CC FT /tag= c
CC FT /trans_except= codon: FAG, a.a.:GU
CC
CC 307041560-1
CC DP 10-SEP-1994
CC DP 23-SEP-1994
CC DP 29-JUL-1993; 188147.
CC PA (MITU) MITSUBISHI KASEI CORP.
CC CYCLUSINULOOIGOSACCHARIDE FRUCTANOTRANSFERASE AND ITS DNA-
CC CYCLUSINULOOIGOSACCHARIDE FRUCTANOTRANSFERASE AND ITS DNA-
CC Claim 2.1; Figure 13-15; 16pp; Japanese.
CC Claim 5.1; Figure 2; 1 page.
CC This nucleotide sequence of the novel cyclinulooligosaccharide
CC fructanotransferase was isolated by PCR using the primers 081186-7 which were
CC

```

Clin 13 J; Flg 11: lllpp; English
This sequence represents the coding sequence for the Drosoephia melanogaster TIM protein. The TIM protein functions as a co-repressor of transcription factors (TF) nuclear translocation protein (NTP) of a TF complex. The NTPs of the invention have specific binding activity to a protein involved in circadian rhythms, and cyclic transcription patterns related to the sleep-wake cycle. The NTP is also preferably light sensitive. The NTP also has the ability to aid the process of circadian rhythm entrainment to environmental cycles of light. TIM has specific binding activity for the Drosoephila period (PER) gene. PER is a nuclear protein which has homology to the family of transcription factors containing the PAS domain. The function of PER is unknown. The amount of PER fluctuates with a circadian rhythm. The NTPs of the invention, their fragments, agonists, antagonists or derivatives are useful for treating depression, narcolepsy or jet lag. The products can also be used for detecting and/or measuring conditions so as to classify groups of individuals with disorders.

Sequence 5198 BP: 1423 A: 1356 C: 1273 G: 1145 T:
Query Match 1.3% Score 14: DB 39: Length 5198;
Best Local Similarity 100.0%; Pred.No. 3.47e+02;
Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Ddb 1132 ttcacagagctcag 1145
|||||||
QY 84 TTCACAGGCTCAG 97

RESULT 526
IDB Q74657 standard; DNA: 5238 BP.
CC Q74657;
AT CC Q74657;
DT 24-JUN-1995 (first entry)
KW LMBI; repetitive element; xerops brassica napus; plant;
KW blackleg infestation; polymerase chain reaction; fungus; ds.
CN Leptospheeria maculans.
PS P0421788.A.
PR 14-MAR-1994; CAO130.
PP 22-MAR-1993; CAO92115.
PA (TAYL./ TAYLOR J L.
PA CANA.) NAT RES COUNCIL CANADA.
DR WPI: 94-317009/29.
PT Testing for infection of plant tissue with Leptospheeria
PT maculans - by PCR using primers derived from the LMBI repetitive
PT elements of L. maculans specific
PT Cloning of L. maculans specific
PT strains

US-08-487-977-3.rgs

generated based on the amino acid sequence of the purified protein (R74634). The resulting fragment was used as probe to obtain the complete DNA sequence (G88185). The protein or bacteria producing the protein, can be used to produce cyclic inulo-oligosaccharide fructanotransferase from beta-2,1-fructose oligosaccharide, 1-fructose oligosaccharide, and alpha-D-glucose.

Digestion conditions:
Restriction Enzyme Incubation Temp Time
SmaI 37°C 3 hr

Sequence 5248 BP: 1370 A: 1212 C: 1502 G: 1163 T:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,476+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB	184 ttggsggttgagctt 197
Bbbs	
Oy	616 ttgsgccttcacct 629

RESULT 528

ID	T80159 standard; cDNA; 5285 BP.
CC	DT0159
CD	T80159.1998 (first entry)
DE	K01591 Inositol 3-kinase cd-mRNA.
DK	Phosphatidylinositol 3-kinase; signal transduction; cell cycle;
DZ	phosphoinositide 3-OH kinase; phosphatidyl transferase; cancer;
KW	oncogene; inflammatory cell disease; cell proliferation; cancer;
RN	pneumonia; retinosis; arteriosclerosis; therapy; diagnosis; mouse;
KS	Mus musculus.
Key	Location/Qualifiers
FF	CDS
FT	3,518
PT	/tag= a
PF	W07J1L650-A1
PD	04-SEP-1997.
PP	12-FEB-1997; U02193.
PB	29-FEB-1996; US-609049.
PR	REGC 3 UNIV CALIFORNIA,
PS	Genbank, California LT:
DR	WP1: 37-448442/A1.
DD	P-P5D8: W38756.
DC	New isolated phosphatidylinositol-3 kinase polypeptide - used to
DT	develop products for diagnostic and therapy, particularly for
DI	cancer metastasis and inflammatory joint diseases, or cancer
DP	Glaun 17; Fig 10; 77pp; English.
PE	This cDNA sequence codes for cpx-m (see W38756), a murine
PC	polypeptide that belongs to a novel class of phosphatidylinositol
CC	kinases that contain a C2 domain are capable of phosphorylating
CC	a diverse range of substrates including phosphatidylethanolamine
CC	phosphatidylinositol 4-phosphate, but not in phosphatidylinositol
CC	4,5-bisphosphate, and which are involved in cell signalling cascades
CC	that control e.g. cell cycle progression and intracellular protein
CC	trafficking. This gene encodes a 518 aa protein derived from a murine
CC	cDNA library (accession #U02193) constructed from a murine
CC	tumor tissue sample (accession #M2349).

26-JUN-1996 (first entry)
DE A: 11199 A; 1348 G; 1478 T; Length 5643;
KW area gene: protease-free; activator: host cell; ss.
OS Aspergillus oryzae strain IPO4177.
FH Key Location/Qualifiers
FT exon 1..2700
FT /tag= a
FT /codon_start= 2282..2284
FT intron 2701..2769
FT /tag= b
FT exon 2770..5643
FT /tag= c
PN WO9535385-A1.
PD 28-DEC-1995.
PF 19-JUN-1995; DKO254
PI (NOVO) NOVO-NORDISK AS.
PI Christensen T, Rynes MJ;
DR WP1: 96-058422/05.
DR P-PSDB: R88408.
PT susceptible to degradation by protease(s)
PT Claim 22: Page 18-22; 42pp; English.
CC The area gene (T09946) of Aspergillus oryzae IPO4177 codes for
CC regulatory area activator protein (R88404), that controls nitrogen
CC metabolism in the fungus. The gene was cloned by cross-hybridization to the
CC A. nidulans area gene. Inactivation of the area gene, by deletion
CC or by use of antisense technology, yields protease-free Aspergillus
CC cells, e.g. strains capable of growing on hosts for prodn. of
CC industrial or therapeutic proteins, such as enzymes and
CC recombinant proteins; ss.
SQ Sequence 5643 BP; 1199 A; 1348 G; 1478 T;
Query Match 1.38; Score 14; DB 18; Length 5643;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 910 tgaacgcatctctgc 923
QY 824 TGACCGCATCTCTGC 837

RESULT 538
AC T97394.
DE 11-MAY-1998 (first entry)
DE Aspergillus oryzae area regulator gene.
KW Area regulator: pepc; pepc: protease; gene replacement;
KW recombinant protein; ss.

OS Aspergillus oryzae strain IPO 4178.
FH Key Location/Qualifiers
FT exon 1..2700
FT /tag= a
FT /note= "contains 1 intron"
FT intron 2701..2769
FT /tag= b
FT exon 2770..4931
FT /number= 1
FT /tag= d
FT /number= 2
PN WO972705-A1.
PD 26-JUN-1997.
PF 15-DEC-1995; DKO528
PI (NOVO) NOVO-NORDISK AS.
PI Christensen T, Lehmbeck J;
DR WP1: 97-341596/31.
PT New fungus lacking functional area and extracellular protease
PT gene(s) - for high yield expression of industrial or therapeutic
PT proteins, also new protease(s) from Aspergillus oryzae
PT (see also WO972705-A1).
CC This DNA (see WO972705-A1) codes for the area regulator
CC protein (see W16130) of Aspergillus oryzae. The area gene
CC was cloned by cross-hybridization with the Aspergillus niger area
CC gene. The invention relates to a new fungus in which: (a) the area
CC gene is inactivated by recombinant DNA methods so that it cannot
CC express a functional area protein; (b) the area gene is inactivated
CC by protease pepc (see T97393) and/or pepc (see T97392) genes are
CC inactivated so that functional enzyme is not produced. The
CC fungi have complete or partial deletions of the appropriate genes,
CC and the genes are inactivated by antisense mechanisms or by
CC replacement. The new fungi are used for the recombinant production
CC of industrial or therapeutic peptides and proteins, particularly
CC enzymes (e.g. proteases, lipase, cellulase or chymosin),
CC tissue plasminogen activator, erythropoietin or thrombopoietin,
CC platelet-derived growth factor, factor VII, factor VIII, urokinase,
CC tissue plasminogen activator, erythropoietin or thrombopoietin.
CC The new fungi do not produce protease and so provide higher yields
CC of recombinant protein because of reduced proteolytic degradation.
SQ Sequence 5043 BP; 1199 A; 1348 G; 1478 T;
Query Match 1.38; Score 14; DB 38; Length 5643;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

910 tgaacgcatctctgc 923
824 TGACCGCATCTCTGC 837
RESULT 539
ID T41406 standard; cDNA to mRNA; 5648 BP.
DE 13-NOV-1997 (first entry)
DE Human REST protein cDNA clone.
DE REST protein: REST-silencing transcription factor; neuron;
KW neural cell; differentiation; neurodegenerative disease;
OS gene therapy; brain cancer; ds.
FH Key Location/Qualifiers
FT cds 325..3618
FT /tag= a REST protein
FT /tag= b complement (805..828)
FT primer_bind
FT /tag= b
FT /note= "5' primer for cDNA amplification"
FT primer_bind 1411..1434
FT /tag= d
FT /note= "3' primer for cDNA amplification"
FT repeat_unit 2284..2241
FT /tag= d
FT repeat_unit 2395..2442
FT /tag= f
FT repeat_unit 2443..2490
FT /tag= f
FT repeat_unit 2491..2628
FT /tag= g
FT repeat_unit 2629..2677
FT /tag= h
FT repeat_unit 2678..2724
FT /tag= i
FT misc_difference 4768
FT /tag= j
FT /note= "base 4768 is given as n in the
FT misc_difference 4893
FT /tag= k
FT /note= "base 4893 is given as n in the
FT /tag= l
FT /note= "specification".
PN WO9629433-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; DKO390823
PI (UNIV) UNIV NEW YORK STATE RES FOUND.
PI Chong JA, Mandel G;
DR WP1: 96-443206/44.

P-PSDB: R93365.
PT New REST-silencing transcription factor (REST protein) - used
PT therapeutically to control differentiation and activity of neural
PT cells, e.g. in case of brain cancer
PT This cDNA clone (T41406) codes for human REST (R93365), or REST-
PT silencing transcription factor, a protein that inhibits the
PT expression of neural proteins in non-neural tissues by silencing
PT expression of neural proteins (see also T41407). The clone was
PT obtained from neural cells of the cerebellum of a human fetus.
PT sequences (see also T41409-11) and subsequent cloning of 2
PT overlapping sequences (T41412-13) and subsequent cloning of 2
PT to produce REST, or REST DNA binding domain (see also R93364), in
PT conditions associated with severe mental retardation and autistic
PT conditions can be used to treat disorders involving de-differentiated
PT neural cells.
SQ Sequence 5648 BP; 1747 A; 1098 C; 1294 G; 1507 T;
Query Match 1.38; Score 14; DB 24; Length 5648;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3003 tcttcacaaactag 3015
Cp 961 TCTTCACAAACTAG 948

RESULT 540
ID Q13572 standard; DNA; 5706 BP.
AC Q13572;
DE 29-NOV-1991 (first entry)
DE T7SA transcription unit.
DE T7SA transcription unit: T7SA transcription unit; hormone, ss.
OS Drosophila melanogaster.
FH Key Location/Qualifiers
FT cds 691..4404
FT /tag= a
FT /tag= b
FT /note= "In cDNA clone Dm4925 T-C."
FT /tag= c
FT /tag= d
FT /tag= e
FT /tag= f
FT /tag= g
FT /tag= h
FT /tag= i
FT /tag= j
FT /tag= k
FT /tag= l
FT /tag= m
FT /tag= n
FT /tag= o
FT /tag= p
FT /tag= q
FT /tag= r
FT /tag= s
FT /tag= t
FT /tag= u
FT /tag= v
FT /tag= w
FT /tag= x
FT /tag= y
FT /tag= z

	P	Plasmid pMR3169 is a tandem plasmid encoding both subunits, i.e.
	p35	(see M4004) and p40 (see W4005), of murine interleukin-12
	IL-12). Each subunit gene was cloned from a mouse spleen cDNA	
	library and has been placed under the transcriptional control of	
	a CMV promoter. The backbone of each subunit contains a strong	
	donor/splicing acceptor is provided between each subuniting	
	its CMV promoter. The backbone of the plasmid is from pcR19,	
	pcR3139 induces at least twice the expression of IL-12 as the	
	bacteriophage vector pHK3136 (see V07043) in vivo and in vitro	
	copies of an expressible foreign genetic construct, especially	
	copies of pMR3169 comprising a promoter operative in the	
	mammalian epidermal cells and DNA sequences encoding p35 and p40	
	construct allows IL-12 expression for treatment of solid tumor	
	metastatic or disseminated tumours, and regression of established	
	tumour. The treatment is effective even when the genetic	
	construct is injected intratumorally after primary tumor; 1766 T;	
	Sequence length inferred 1798 A., 1877 C., 1840 G.; 1766 T; Query Match 1.3%; Score 14.: DB 39.; Length 7287;	
	Best Local Similarity 100.0%. Pred. No. 3.47e+02; Indels 0; Gaps:	
	Gaps 14.; Conservative 0; Mismatches 0;	
Df	1126 ttcttcgactgcgtc 1139 QY 889 TTCCTCAGTGCATGTC 902	
RESULT 558		
ID	T74883 standard; cDNA; 7393 BP. ID T74883 D 09-FEB-1998 (first entry) DE Porcine retrovirus cDNA (defective).	
KW	Retrovirus; porcine; GAG protein; POL protein; ENV proteins; KW activated virus PCRs, infectious; provirus; organ transplant; donor; OS Porcine retrovirus.	
FPE	Key Location/Qualifiers	
CDS	598..2172	
FT	/tag= a b	
FT	/tag= b	
FT	/note= putative GAG protein"	
CDG	/tag= c	
FT	/note= Putative POL coding region (partial)	
mat_peptide	2120..3522	
FT	/note= Putative POL protein (partial)"	

```

RESULT 559
ID N20042 standard; cDNA; 7440 BP.
DT 05-AUG-1992 (first entry)
DE Sequence of a full-length cDNA copy of the poliovirus genome in
DE plasmid pVR106.
OS Homo sapiens
OS Human poliovirus; picornavirus; vaccine; antigen; immunogen; ss.
Key Locaton/Qualifiers
Key 743..949
cda /tag= a
cda /product= P4
cda /tag= b
cda /product= VP2
cda /tag= c
cda /product= VP3
cda /tag= d
cda /product= VP1
cda /tag= e
cda /product= 3b
cda /tag= f
cda /product= 5b
cda /tag= g
cda /tag= h
cda /product= 1b
cda /tag= i
cda /product= VPg
cda /tag= j
cda /product= 4(p63)
W08203632-A.
28-OCT-1982.
12-NOV-1981. 320525
12-NOV-1981. US-120525
12-NOV-1981. US-120525
(PASI) MASSACHUSETTS INST TECH.
Baltimore D, Racaniello VR.
Proc Natl Acad Sci USA 82:950592/44 (950592).
Prodn of cDNA representing viral RNA sequences - by

```

PT transfection, insertion into vector and host cell transformation
 PD Example: Table 1, pages 25-31, 50pp; English.
 PC This is a full-length ORF-5 coding region (1000 bp) of the
 CC HB101 contig. This plasmid has been registered as ATCC 31844. The
 CC full-length poliovirus cDNA molecule is itself infectious and can be
 CC introduced into cells and these cultured to produce RNA virus.
 CC The altered material used to infect cells so that attenuation and
 CC RNA is prod. and this used to make vaccines. For antibody prodn.
 CC cDNA capable of directing antigen prodn. is selected and isolated and
 CC incorporated into cells which are incubated to produce RNA antigen.
 SQ Sequence 7448 BP; 2205 A; 1754 C; 1716 G; 1785 T;

Query Match 1.3%; Score 14; DB 4; Length 7440;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4901 tgaacatgcat 4914
 QY 229 ttaacatgcat 242

RESULT 560
 ID M9312 standard; DNA: 7499 BP.
 AC 19-APR-1989; 17-020314.
 DE PC101 plasmid.
 DE PC101 plasmid.
 DE Probe: hybridisation.
 OS Chlamydia trachomatis.
 PT Key Location/Qualifiers
 FT CDS
 FT /tag= a
 FT /note= orf8 (read 3'-5')
 FT cda 1179..2532
 FT /tag= orf1
 FT cda 2528..3590
 FT /tag= c
 FT /note= orf2
 FT cda 3584..4416
 FT /tag= d
 FT /note= orf3
 FT cda 4518..4824
 FT /tag= e
 FT cda 4917..5646
 FT /tag= f
 FT /note= orf5
 FT cda 5645..6388
 FT /tag= g

PT cda /note= orf6
 FT 6620..7402
 FT /tag= h
 FT /note= orf7
 PD EP-336412-A.
 PC 06-APR-1989; 17-020314.
 CC 08-APR-1989; 17-020314.
 CC Ratti G, Conanducci M, Ricci S, Garuti G, Cosco E;
 DR WPI: 89-28424/41.
 PT Synthetic oligonucleotides specific to Chlamydia trachomatis +
 PT plasmid for detection of C. trachomatis in biological samples
 CC PC101 can be isolated from C. trachomatis L2/434/Bu (ATCC VR 902) grown
 CC on HeLa cells. Synthetic oligonucleotides corresp. to 10 or more
 CC nucleotides of the plasmid, were used for detection kits, esp. oligos
 CC corresp. to 400 BP between ORF6 and ORF7.
 SQ Sequence 7499 BP; 2460 A; 1286 C; 1432 G; 2331 T;

Query Match 1.3%; Score 14; DB 1; Length 7499;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7194 ataaagaaccacaa 7207
 Cp 647 ATAAAGAACCCAA 634

RESULT 561
 ID Q27429 standard; DNA: 7502 BP.
 AC 02-APR-1993 (first entry)
 DE PCTD.
 DE CT: PCTD: epithelium; ocular mucosa; uro-genital mucosa; antigen;
 KW Monoclonal; polyclonal; antibody; vaccine; ss.
 OS Chlamydia trachomatis
 PT Key Location/Qualifiers
 FT CDS
 FT /tag= a
 FT /label= ORF-1
 FT cda 1480..3592
 FT /tag= b
 FT /label= ORF-2
 FT cda 3604..4398
 FT /tag= c
 FT /label= ORF-3
 FT cda 4468..4776
 FT /tag= d
 FT /label= ORF-4
 FT cda 4804..5598

PT cda /tag= e
 FT /label= ORF-5
 FT 5595..6338
 FT /tag= f
 FT /label= ORF-6
 FT cda 6560..7489
 FT /tag= g
 FT /label= ORF-7
 FT cda 7502..8242
 FT /tag= h
 FT /label= ORF-8

EP-499681-A.
 PD 15-AUG-1992; 106310.
 PR 07-FEB-1991; 17-020314.
 PA (ISTS) SCLAVO SPA.
 PI Conanducci M, Giuliani M, Ratti G, Tecce MP;
 DR P-RSD: 82961/35.
 PT PCTD plasmid from Chlamydia trachomatis and immunogenic proteins
 PT - for diagnosing and vaccinating against Chlamydia infections
 PT e.g. venereal lymphogranuloma
 OS Chlamydia trachomatis
 CC The plasmid is 16.4 kbp; English
 CC (CT) serotype D. PCTD. This serotype generally infects epithelial
 CC tissues, such as the ocular and uro-genital mucous membranes, and
 CC shows a low virulence. The isolated plasmid encodes eight proteins,
 CC the sense strand and the eight on the complementary strand.
 CC Preparation of poly- and mono-clonal antibodies to be used for the
 CC diagnostics. The antigens can also be used in the formulation of
 CC vaccines against infections due to CT.
 SQ Sequence 7502 BP; 2460 A; 1286 C; 1430 G; 2324 T;

Query Match 1.3%; Score 14; DB 4; Length 7502;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7146 ataaagaaccacaa 7159
 Cp 647 ATAAAGAACCCAA 634

RESULT 562
 ID T85043 standard; DNA: 7766 BP.
 AC 28-04-1989 (first entry)
 DE Borrelia variabilis (WPI) -like sequence (vls) locus.
 KW Variable major protein-like sequence; VMP-like sequence; vls locus;
 KW vlsE gene; Lyme disease; relapsing fever; therapy; diagnosis;
 KW vaccine; ss.

OS Borrelia burgdorferi strain B31-543 (ATCC 35210).
 Key Location/Qualifiers
 FT misc_feature 205..711

FT /tag= a
 FT /label= vls2
 FT misc_feature 712..1293

FT /tag= b
 FT /label= vls3
 FT misc_feature 1294..1869

FT /tag= c
 FT /label= vls4
 FT misc_feature 1870..2439

FT /tag= d
 FT /label= vls5
 FT misc_feature 2440..3009

FT /tag= e
 FT /label= vls6
 FT misc_feature 3010..3493

FT /tag= f
 FT /label= vls7
 FT misc_feature 3484..3990

FT /tag= g
 FT /label= vls8
 FT misc_feature 3991..4516

FT /tag= h
 FT /label= vls9
 FT misc_feature 4517..5038

FT /tag= i
 FT /label= vls10
 FT misc_feature 5039..5652

FT /tag= j
 FT /label= vls11
 FT misc_feature 5653..6219

FT /tag= k
 FT /label= vls12
 FT misc_feature 6220..6789

FT /tag= l
 FT /label= vls13
 FT misc_feature 6790..7373

FT /tag= m
 FT /label= vls14
 FT misc_feature 7374..7946

FT /tag= n
 FT /label= vls15
 FT misc_feature 7947..8000

FT /note= 'the vls sequence provided is only
 FT /note= 'the vls sequence provided is only

7766 bases"

W97371123-AL.
28-AUG-1997. H02952.
21-SEP-1996; US-101208.
(TEXT) UNIV TEXAS SYSTEX.
Barbour AC, Hardham JM, Howell JK, Morris SJ, Weinstein GM;
UR 97-435172-40.
Nucleic acid encoding variable major protein-like peptide of
Borrelia - useful for recombinant production of Vsp like protein or
antigen. *Journal of Molecular Biology*. 1997; 270: 1-12.
Claim 26. Page 101-103, 110pp. English.
This DNA sequence comprises the variable major protein (VMP)-like
sequence (vls) locus of *Borrelia burgdorferi*. An infectivity
assay was performed using the recombinant protein (Vsp-like protein)
isolated by subtractive hybridization. It contained the vls locus,
which consists of 15 silent vls cassettes (vls1-vls16) and the 15
expressed vls gene (see 785043). Portions of several of the 15
cassettes were subcloned into the pET3a vector and used to produce
Vsp, recombine into the central Vsp cassette region during
infection, resulting in antigenic variation of the expressed
lipoprotein (see W227676) and hence immune evasion, long-term
infection, and disease. The recombinant protein (Vsp-like protein),
including vls2-vls16 sequences, encoding Vsp-like proteins, can be
used for the recombinant production of Vsp-like proteins, or for
the diagnosis of Lyme disease (claimed), and may also have
other uses (claimed).
Sequence 7766 BP. Accession 2039 A. 867 C. 3011 G. 1896 T.
390

Query Match 1.3%; Score 14; DB 35; Length 7766;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB	3400	gagaagcctgagga	3413
DB			
DB	1013	CGAGCCTGAGGA	1000

RESULT 563	
ID	Q84660 standard: DMA: 7791 bp.
AC	Q84660:
OR	01-052-1995 (first entry)
PR	Human neuronal calcium channel subunit alpha 1a-2.
NA	Human neuronal calcium channel subunit alpha 1a-2.
EW	Lambert Eaton Syndrome: sg.
OS	Homo sapiens.
Key	Location/Qualifiers
cds	237..7037
FT	/*cds= a
FA	W09504822-A.

Tue Nov 17 08:55:26 1998

US-08-887-977-9.IMP

```

PT      misc_difference 7035..7039
PT      /note= "not present in alpha 1A-2"
PT      "acc= b
PT
PT      W09504822.A.
PT      16-FEB-1995.   U09330
PT      11-AUG-1993.   U8-105536.
PT      05-NOV-1993.   US-148097.
PT      (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
PT      Ellis SP, Gillespie A, Harpold WH, McCue AF, Williams ME;
PT      P-SPDS; R71007/12.
PT      P-SPDS; R71007/12.
PT      DNA encoding human calcium channel sub-unit(c) - used for
PT      developing probes. for studying calcium channels, e.g. for
PT      obtaining agonists and antagonists
PT      Tiam-1 transmembrane-107 at the English.
PT      One form, alpha 1A-1 is given in Q84659/R71007, and the other,
PT      alpha 1A-2 is given in Q84660/R71008. Alpha 1A-2 differs from
PT      alpha 1A-1 by 10 amino acids, and in that it lacks a
PT      start-codon-encoding sequence at the first transmembrane
PT      region. This sequence is a start-codon-encoding sequence that
PT      introduces a translation termination codon resulting in an
PT      alpha 1A-2 coding sequence that encodes a shorter alpha 1A
PT      subunit than that encoded by alpha 1A-1. DNA encoding alpha
PT      1A subunits can be obtained using Q84660 or Q84659. The DNA
PT      encoding alpha 1A subunits can be obtained using Q84660 or
PT      the phage lysate of an E. coli host contg. DNA encoding an alpha
PT      1A subunit that has been deposited in the ATCC under accession
PT      no. 75293. The DNA is such a phage includes the DNA fragment
PT      containing the sequence in Q84661 which selectively hybridizes under
PT      conditions of high stringency to DNA encoding alpha 1A but not
PT      to DNA encoding alpha 1B.
PT      Sequence 7808 BP. 1680 A; 2441 C; 2265 G; 1422 T;
PT      90

```

Query Match	1.38;	Score 14;	DB 15;	Length 7808;
Best Local Similarity	100.08;	Pred. No.	3.47e+02.	

Matches 14; Conservativ

QY 342 CATCTATGCCATCA 355

ID AC
T74884
09-FEB-1998 (first entry)
DE Miniature swine retrovirus cDNA
DT
DT
KW Retrovirus; porcine; GAG protein;
KW xenotransplantation; infectious; grovirus; organ transplant; donor;
T74884 standard; cDNA; 7892 BP.

16-FEB-1995
PD 16-DEC-1994; Q09230.
PR 11-DEC-1993; US-1405736.
PR 03-NOV-1991; US-1405736.
PR 03-NOV-1991; US-1405736.
PR Ellis 69, Gillespie A., Harpold MW, McCune AF, Williams ME;
D1 WP1: 95-08000/J2.
D1 WP2: 95-08000/J2.
D1 DNASISB: 871008.
D1 DNASIS: 871008.
PT developing probe for studying calcium sub-unit(s) - used for
PT obtaining agonists and antagonists
PT obtaining agonists and antagonists
PT Tamm, in Page 190-201; 25pp; English.
PT Tamm, in Page 190-201; 25pp; English.
PT alternatively spliced to yield at least two variant mRNAs.
CC One form, alpha 1, is given in Q84659/R1007, and the other,
CC alpha 1.2 is given in Q84659/R1008. Alpha 1.2 differs from
CC alpha 1.1 in 28 positions. Alpha 1.2 encodes a protein with a
CC 5 nt sequence. This deletion shifts the reading frame and
CC introduces a translation termination codon resulting in an
CC alpha 1.2 subunit that encodes a truncated alpha 1 subunit
CC subunit than that encoded by subunit 1.1. DNA encoding alpha
CC alpha 1 subunits can be isolated using all or a portion of the DNA
CC having sequence Q84651, Q8569 or Q84660 or DNA abcd. From
CC alpha 1 subunit that has been deposited in the ATCC encoding an alpha
CC 1A subunit that has been deposited in the ATCC encoding an alpha
CC no. 75293. The DNA is such a phase includes the DNA fragment
CC having the sequence in Q84661 which selectively hybridizes under
CC standard conditions to DNA encoding alpha 1A DNA but not
CC to DNA encoding alpha 1B.

sq sequence 7791 BP; 1675 A; 2436 C; 2458 G; 1422 T;
 Query Match 1.30; Score 14; DB 15; Length 7791;
 Best Local Similarity 100.00; P-val 0.0 3.47e+02.

Matches 14; conservative

Db 5363 catctatgccatca 53766
|||||

RESULT	564	
ID	Q84659	standard; DNA; 7808 BP.
AC	Q84659;	
DT	11 DEC-1995	(first entry)
DR	Q1 DEC-1995	(second entry)
CH	Calcium channel	antagonist; agonist; diagnosis;
KW	Calcium channel	subunit; antagonist; agonist; diagnosis;
OS	Lambert Eaton Syndrome;	as.
KS	Lambert Eaton Syndrome;	as.
OW	Homo sapiens.	
FT	key	Location/Qualifiers
FR	key	131,7765
FI	cds	4240
FT	cds	4240

Tue Nov 17 08:55:26 1998

US-08-887-977-9.IMP

TV	activated virus; PCR; ss.		
TV	porcine retrovirus	Location/Qualifiers	
FT	Key	585..2159	
CD5		585..2159	
FT	mat_peptide	*tag= b	
FT		/note= "putative GAG protein"	
FT	CD5	3307..5744	
FT		*tag= d	
FT	mat_peptide	3307..5741	
FT		/note= "putative POL protein"	
FT	CD5	5670..7536	
FT		*tag= f	
FT	mat_peptide	5670..7533	
FT		/note= "putative ENV protein"	
W0721086-N1.			
TV	19-721087		
PR	13-DEC-1996; U19680.		
PR	14-DEC-1995; US-577645.		
PA	(GHERO) GEN HOSPITAL CORP.		
DR	WPI397-332804/310.		
DR	P-PSDB: W32095-W32098.		
DR	New nucleic acid from porcine retro-viruses - used for detecting		
DR	viruses in transplant or other tissue and for assessing risk of		
PT	transmission to the recipient		
PT	Claim 23: Pg 3, 18pp English		
CC	This cDNA sequence represents a porcine retrovirus from miniature swine		
CC	containing the putative coding regions for viral GAG, POL and ENV		
CC	protein. This sequence and PCR fragments generated from it		
CC	porcine retroviruses prior to xenotransplantation. Transplantation		
CC	increases the likelihood of retroviral activation if intact and infectious		
CC	proteins are present. The porcine retroviral sequence can be used to		
CC	generate probes to determine the level (e.g. copy number) of intact		
CC	retroviruses in a sample. The sequence can be used to detect mutations,		
CC	xenotransplantation donors. It can be used to detect mutations,		
CC	genetic lesions or viral recombinants and also to determine the		
CC	histological localisation of activated retrovirus. Using Polymerase Chain		
CC	reaction DNA quantitation (PQ) on blood mononuclear cells, infectivity		
CC	assays and immunofluorescence can be used to detect retroviral		
CC	donors without intact porcine retroviral sequences or a lower copy number		
CC	of viral elements could be selected.		
CC	Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T;		
SQ			

Query Match 1.39; Score 14; DB 35; Length 7092;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Total 0;

QY	1062	CCCCCCCCCCCCCCCC	GGCAGATACGACA	1075
		RESULT	574	
ID	IC	Q4241	standard: DNA: 8700 BP.	
AC	Q42541			
DT	15-SEP-1993	(first entry)		
VS	B911/Hpall fragment of pR32.			
PR	Staphylococcus epidermidis			
PF	Key	Location/Qualifiers		
FE	complement (1227..781)			
FT	/*tag= a	/*note= "epiP CDS"		
FT	/*tag= b	/*note= "epiP CDS"		
FT	/*tag= c	/*note= "epiP CDS"		
FT	complement (6983..61866)			
FT	/*tag= b	/*note= "epiQ CDS"		
FT	complement (6379..6994)			
FT	/*tag= c	/*note= "epiP CDS"		
FT	complement (1381..4584)			
FT	/*tag= e	/*note= "epiA CDS"		
FT	complement (4443..5808)			
FT	/*tag= e	/*note= "epiA CDS"		
FT	complement (5824..6359)			
FT	/*tag= f	/*note= "epiD CDS"		
FT	complement (5394..4585)			
FT	/*tag= f	/*note= "epiB CDS"		
EP	543195-A.			
26-MAY-1993.	118598			
31-OCT-1991.	DS-784234			
(THOM) THOMAS GRESH KARL.				
PA	Augustin J, Empecke K, Entian R, Gots P, Jung G, Maletta C, Klein C.			
PI	W1103:1680372,1, Roosenkrantz R, Schnell N, Wieland B;			
PR	PF-D59: R37115, R39343, R39344, R39345, R39352, R39353, R39354.			
DD	Roval DNA molecule - encoding Rpi.B, C, D, P or Q enzymes			
DR	This is the nucleotide sequence of the BglII/HpaII fragment of the			
CC	involved in biosynthesis of lantibiotic epidermin.			
CC	Staphylococcus epidermidis plasmid pR432 which contains open reading			
CC	frames for epiA, epiB, epiC, epiD, epiP, epiQ and epiY. The proteins			
CC	encoded by these are valuable reagents potentially useful in the			
CC	mass-lanthionine, 3-methyl-lanthionine and 5-(2-aminoethyl)-D-cysteine.			
CC	meso-lanthionine, 3-methyl-lanthionine and 5-(2-aminoethyl)-D-cysteine.			

```

FF FF      /label= cpe
FF FF misc_signal 1055..1066
FF FF /*tag= k
FF FF /*label= BP-II-IbctdA
FF FF tata_signal 1073..1078
FF FF /*tag= l
FF FF exon 1102..1170
FF FF /*tag= n
FF FF /*number= I
FF FF repeat_unit 1405..1629
FF FF /*tag= n
FF FF /*rpt_type= DIRECT
FF FF /*note= porcine-specific repetitive element PRF-1"
FF FF exon 1909..1970
FF FF /*tag= o
FF FF /*number= II
FF FF /*rpt_type= DIRECT
FF FF exon 2499..2550
FF FF /*tag= p
FF FF /*number= XII
FF FF exon 3707..3902
FF FF /*tag= q
FF FF /*number= IV
FF FF repeat_unit 4022..4044
FF FF /*tag= r
FF FF /*note= "cag"-like element"
FF FF repeat_unit 4455..4716
FF FF /*tag= s
FF FF /*rpt_type= DIRECT
FF FF exon 5000..5211
FF FF /*tag= t
FF FF /*number= V
FF FF exon 6474..6604
FF FF /*tag= u
FF FF /*number= VI
FF FF exon 7384..7596
FF FF /*tag= v
FF FF /*number= VII
FF FF /*tag= w
FF FF /*tag= w
FF FF /*label= UA-motif
FF FF poly_signal 8164..8270
FF FF /*tag= x
FF FF repeat_unit 8614..8760
FF FF /*tag= y
FF FF /*rpt_type= DIRECT
FF FF /*note= "porcine-specific repetitive element PRF-1"
FF FF NE9301929-A
FF FF 17-OCT-1994.
PD PD

```

FT mat_peptide 4587..5310
 FT /tag= f
 FT /product= green fluorescent protein
 FT 280..349
 FT /tag= g
 FT /product= Shiva-1
 PR W09640879-A1.
 PD 09-DEC-1995. UN0041.
 PR 07-JUN-1995. US-480454.
 PA (UTAH) UNIV UTAH STATE.
 PI Morrey, J., Reed WA, White KL;
 PR 07-DEC-1995. US-480454.
 PT Expression of disease-resistant animals or for the treatment of diseases
 PT prodn. of disease-resistant animals or for the treatment of diseases
 PT e.g. tumours
 Claim 11: Page 43-48; 75pp; English.
 CC The present invention relates to express the amphipathic (lytic)
 CC peptide, Shiva-1, in the milk of transgenic animal. The Shiva-1
 CC gene is placed under control of bovine milk-specific beta casein 5'
 CC regulatory elements. The cassette also encodes a green fluorescent
 CC protein that acts to inhibit the toxic activity of the amphipathic
 CC peptide. The present invention also includes a method for isolating and
 CC purification. Another cassette (see also T08055) also includes
 CC beta casein 3' regulatory elements. Other, interleukin-regulated
 CC cassettes can be used to produce disease-resistant animals or to
 CC treat diseases, e.g. tumours.
 SQ Sequence 8197 BP; 2874 A; 1814 C; 1806 G; 2503 T;
 Query Match 1.38; Score 14; DB 27; Length 8797;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 3689 tcttggtttcttt 3702
 QY 632 TCTTGTCTTCTT 645
 RESULT 577
 ID T03104 standard; DNA: 9108 BP.
 AC T03104.
 DE Plasmid pRK5 tk1-1.1 encoding Sal S-1.
 KW Protein tyrosine-kinase; pTK; SAL-S1; agonist; cell growth;
 KW differentiation; pRK5 tk1-1.1; ss.
 OS Chimeric synthetic;
 OS Chimeric synthetic;
 PR W09527051-A1. sepians.
 PD 12-OCT-1995.
 PF 04-APR-1995; 004228.
 PR 04-APR-1994; 09-222616.

CC (see also Q64176) from a heterogenic virus as the the NANRV protein
 CC may be expressed by a transformed insect cell. The expressed protein
 CC These antibodies may be used in the generation of anti-NANRV antibodies.
 SQ Sequence 9391 BP; 1862 A; 2836 C; 2698 G; 1995 T;
 Query Match 1.38; Score 14; DB 11; Length 9391;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 8533 tgccttcttcgcg 8546
 Cp 1075 TGTCTTATCTCG 1062
 RESULT 579
 ID Q38559 standard; cDNA: 9391 BP.
 AC Q38559.
 DE 9-JUL-1993 (first entry)
 KW HCV; detection; antigen; vaccine; recombinant; ss.
 OS Hepatitis C virus.
 PR Key
 FT Location/Qualifiers
 FT CDS
 PR J05068563-A.
 PD 23-MAR-1993.
 PF 17-JUL-1991; 201884.
 PA (KAGAKU) KAGAKU OTSUKI KESSEI NYOHO.
 DR WPI: 93-130639/16.
 DR P-PDB: R35207.
 PT Nucleotide sequence encoding hepatitis C virus polypeptide - is
 PT against HCV
 Claim 1: Page 6-17; 17pp; Japanese.
 CC RNA was extracted from the plasma of Japanese patients whose HBs
 CC antigen was negative and with a GGT over 100. cDNA was synthesised
 CC with RNAse inhibitor and a GGT over 100. cDNA was synthesised
 CC infected chimpanzee plasma to isolate HCV clones. The DNA sequence
 CC is useful in detection of HCV virus. The polypeptide it produces
 CC may be used as an antigen in the prep. of HCV vaccine. 1994 T;
 SQ Sequence 9391 BP; 1861 A; 2836 C; 2698 G; 1995 T;
 Query Match 1.38; Score 14; DB 6; Length 9391;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 8533 tgccttcttcgcg 8546
 Cp 1075 TGTCTTATCTCG 1062

PP 05-NOV-1993; 001929.
 PR (REDA) 36160/47.
 PI Billiau AJDA, Vandenbroeck K;
 DR WPI: 94-322838/40.
 DR P-PDB: R63137.
 PT Recombinant human IL-1 beta and new recombinant
 PT Immunostimulant
 Claim 3: Fig 2; 40pp; Dutch.
 CC The genomic DNA sequence coding for porcine interleukin-1 beta was
 CC as a metabolic growth factor and lipid metabolism. It is also useful
 CC as an immunostimulant, esp. as a vaccine adjuvant.
 SQ Sequence 8760 BP; 2318 A; 2087 C; 1989 G; 2365 T;
 Query Match 1.38; Score 15; DB 12; Length 8760;
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 5269 tcttggtttcttt 5283
 QY 146 TCTTGTCTTCTT 160
 RESULT 576
 ID T50856 standard; DNA: 8797 BP.
 AC T50856.
 DE Amphipathic peptide production in milk
 KW Shiva-1; lytic peptide; DNA cassette; disease resistance;
 KW gene therapy; tumour; ds.
 OS Chimeric Bos taurus;
 OS Chimeric synthetic.
 PR Key
 FT Location/Qualifiers
 FT 5'utr
 FT 127..1800
 FT /tag= b
 FT tatta_signal
 FT 1769..1773
 FT /tag= b
 FT exon
 FT 1801..1823
 FT /tag= c
 FT /tag= c
 FT 1987..1992
 FT /tag= d
 FT /note= "beta casein exon 2"
 FT exon
 FT 4567..4590
 FT /tag= b
 FT /note= "Portion of beta casein exon 3"

PA (GETH) GENENTECH INC.
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
 PI Wood WJ;
 PR (REDA) 36160/47.
 PI Billiau AJDA, Vandenbroeck K;
 DR WPI: 94-322838/40.
 DR P-PDB: R63137.
 PT Recombinant human IL-1 beta and new recombinant
 PT Immunostimulant
 Claim 3: Fig 2; 40pp; Dutch.
 CC The genomic DNA sequence coding for porcine interleukin-1 beta was
 CC as a metabolic growth factor and lipid metabolism. It is also useful
 CC as an immunostimulant, esp. as a vaccine adjuvant.
 SQ Sequence 8760 BP; 2318 A; 2087 C; 1989 G; 2365 T;
 Query Match 1.38; Score 14; DB 16; Length 9108;
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4983 cagagagagagag 4996
 Cp 160 CCAGAGAGAGAG 147
 RESULT 578
 ID Q64175 standard; cDNA: 9391 BP.
 AC Q64175.
 DE 09-FEB-1995 (first entry)
 KW NANRV EL/22 gene region.
 KW hcv/22 protein; non-A, non-B hepatitis virus; NANRV; signal peptide;
 KW antibody; hepatitis C virus; HCV; vaccine; ds.
 OS Non-A, non-B hepatitis virus.
 PR Key
 FT Location/Qualifiers
 FT CDS
 PR J06141873-A.
 PD 14-MAR-1994. 089371.
 PA (KAGAKU) KAGAKU OTSUKI KESSEI NYOHO KENKYUSHO.
 DR WPI: 94-205030/25.
 DR P-PDB: R54099.
 PT Nucleotide sequence encoding hepatitis C virus and signal sequence - useful
 PT in vaccines and in immunological detection
 Disclosure: Page 7-19; 23pp; Japanese.
 CC This sequence encodes the EL/22 protein from non-A, non-B hepatitis
 CC virus (NANRV). This sequence may be linked to a signal peptide

PT Human breast cancer susceptibility gene BRCA2 - useful for
PT diagnosing breast cancer and screening for compounds to treat breast
PS Claim 2: Pages 90-106; 18pp; English.
CC The present sequence is the human breast cancer susceptibility gene
CC BRCA2, which can be used to diagnose breast cancer and screen for
CC compounds to treat breast cancer. BRCA2 can also be used in gene
CC construction, which can be used to produce transgenic plants which
CC has lost its or has altered (i.e. by virtue of a mutation in BRCA2)
CC BRCA2 gene function.
SQ Sequence 11385 BP: 4107 A; 1988 C; 2112 G; 3178 T;
Query Match 1.33; Score 14; DB 36; Length 11385;
Best Local Similarity 100.0%; Pred. No. 3.47e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7920 tgaagattatttc 7933

Qy 17 TGAAGATTATTTC 30

RESULT 591
ID T79805 standard; DNA; 11478 BP.

AC T79805;
DT 11-NOV-1997 (first entry)
KW Full length potato starch branching enzyme gene.
KW Transgenic plant; starch production; modification; ss.
OS Solanum tuberosum.
FN Key Location/Qualifiers
FT Promoter 1..2131
FT tata_signal 2048..2051
FT cds 2132..9639
FT exon 2132..2209
FT intron 2132..2209
FT intron 2210..3376
FT exon 3377..3496
FT intron 3497..3817
FT exon 3818..4037
FT intron 4038..4541
FT exon 4542..4611
FT intron 4612..4757
FT exon 4758..5027
FT intron 5028..5245
FT exon 5246..6152
FT intron 6153..6350
FT exon 6351..6467
FT intron 6468..6675
FT exon 6676..6738
FT intron 6739..7031
FT exon 7032..7139
FT intron 7140..7516
FT exon 7517..7617
FT intron 7618..7789
FT exon 7790..7857
FT intron 7858..8002
FT exon 8003..8084

FT exon 8003..8084
FT intron 8085..8126
FT exon 8127..8443
FT intron 8444..9240
FT exon 9241..9636
FT intron 9637..11478
FT 3'utr
FT 8007A119.12.
FT 8007A119.12.
FT 12-JUL-1996; E03052.
FT 14-JUL-1995; GB-014435.
FT (DANF-) DANISCO AS.
FT 8007A119.12.
FT Affecting enzymatic activity in plant by expressing anti-sense
FT intron sequence - esp. for inhibiting starch branching enzyme thus
FT creating modified starch without post-harvest derivatization
FT Promoter; tgaagattatttc 7933; English; potato starch-branching enzyme
CC T79805
CC (SBE) gene. Sequences atttgc to introns of the SBE gene are used
CC to produce transgenic plants producing modified or new forms of starch
CC for industrial use. The sequences allow starches to be made without
CC the use of harsh chemicals and require much less energy.
CC efficient and selective than known methods for regulating enzymatic
CC activity because there is almost no homology between introns in
CC different SBE genes. The SBE gene promoter (see T79804) is also
CC present in the transgenic plants.
SQ Sequence 11478 BP: 3618 A; 1863 C; 2127 G; 3864 T;
Query Match 1.33; Score 14; DB 33; Length 11478;
Best Local Similarity 100.0%; Pred. No. 3.47e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6829 tgaagattatttc 6842

Cp 661 TGAAGATTATTTC 648

RESULT 592

ID T79784 standard; DNA; 11478 BP.

FT intron 4038..4541
FT exon 4542..4611
FT intron 4612..4757
FT exon 4758..5027
FT intron 5028..5245
FT exon 5246..6152
FT intron 6153..6350
FT exon 6351..6467
FT intron 6468..6675
FT exon 6676..6738
FT intron 6739..7031
FT exon 7032..7139
FT intron 7140..7516
FT exon 7517..7617
FT intron 7618..7789
FT exon 7790..7857
FT intron 7858..8002
FT exon 8003..8084

AC T79784;
DT 10-NOV-1997 (first entry)
KW Starch branching enzyme; SBE; potato; genetic engineering;
KW Transgenic plant; starch production; modification; ss.
OS Solanum tuberosum.
FN Key Location/Qualifiers
FT Promoter 1..2131
FT tata_signal 2048..2051
FT cds 2132..9639
FT exon 2132..2209
FT intron 2210..3376
FT exon 3377..3496
FT intron 3497..3817
FT exon 3818..4037
FT intron 4038..4541
FT exon 4542..4611
FT intron 4612..4757
FT exon 4758..5027
FT intron 5028..5245
FT exon 5246..6152
FT intron 6153..6350
FT exon 6351..6467

QY |||||
603 GAGCTGTCGATGGT 617

RESULT 588
D TCCTTTT standard; DNA: 20710 BP.
AD T30681
DE T30681
DT 11-SEP-1996 (first entry)
DD Kaposi's sarcoma associated herpesvirus clone KSS.
DR Kaposi's sarcoma associated herpesvirus; KSHV; gamma 2 herpesvirus;
KW Kaposi's sarcoma associated herpesvirus vaccine; AIDS as.
QS Kaposi's sarcoma associated herpesvirus vaccine; AIDS as.

```

DE Kapozi's sarcoma associated herpesvirus clone K5S.
KW Kapozi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;
OS Kaposi's sarcoma associated herpesvirus; vaccine; AIDS; as.
FH Key Location/Qualifiers
FT cds complement (1..364)
FT /cds= "incomplete ORF20"
FT /tag= b
FT /product= thymidine kinase
FT /note= "ORF20"
FT /tag= c
FT /product= glycoprotein H
FT /note= "ORF22"
FT /tag= d
FT /product= complement (4282..5496)
FT /note= "ORF23"
FT /cds= complement (5500..7758)
FT /tag= e
FT /product= "ORF24"
FT /tag= f
FT /note= "ORF25"
FT /tag= g
FT /product= virion polypeptide VP23
FT /note= "ORF26"
FT /cds= 12851..13723
FT /tag= h
FT /product= minor capsid protein
FT /note= "ORF27"
FT /tag= i
FT /product= "ORF28"
FT /tag= j
FT /product= complement (14343..15485)
FT /tag= k
FT /product= (with ORF29A) putative DNA packaging Protein "ORF29B"
FT /tag= l
FT /cds= 15609..15842
FT /tag= m

```

```

FT      /tag= j
FT      /product= (with OR29A) putative DNA packaging
FT      protein
FT      /note= "OR298"
FT      15609..15842
FT      /tag= k
FT      cds

```

```

Db      7069  cagactctctatctt 7083
      |||||
QY      603  gacactctgacttt 617

RESULT 509
ID      T31137 standard; DNA; 37895 BP.
AC      T31137.
DT      30-JUN-1997 (first entry)
DE      Gene for biosynthetic enzyme group of isopentanoic acid synthase
DI      Isopentanoic acid synthase; IPA; drugs; agrochemicals;
DI      isopentanoic acid synthase; IPA; drugs; agrochemicals;
DI      anti-inflammatory; anticancer agent; ss.
OS      Shewanella putrefaciens.
FH      Key
FH      Location/Qualifiers

```

FF	cds	/product=	Icosapentaenoic acid biosynthetic enzyme
FF		/tag=	9681.12590
FF		/tag=	c
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	1300.13003
FF		/tag=	d
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	1306.22173
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	21176.24515
FF		/tag=	f
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	21176.24525
FF		/tag=	g
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	30730.32358
FF		/tag=	h
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF	cds	/tag=	32753.34327
FF		/tag=	i
FF		/product=	Icosapentaenoic acid biosynthetic enzyme
FF		/tag=	j

(SAGA) SAGAMI CHEM RES CENTRE.
PA Kato S, Kondo K, Yamada A, Yazawa K;
PI WPI: 96-342288/34.
DR P-PSDB: R39458, R39459, R39460, R39461, R39462, R39463, R39464,
DR

RESULT 604
ID T18551 standard; DNA: 53577 BP.
AC T18551:
DE 06-MAY-1997 (first entry)
DT Human polycystic kidney disease normal PKD1 gene.
DR Adult onset polycystic kidney disease; APD; autosomal dominant;
RS mutation: transposition; deletion; insertion. 4

Key	Location/Qualifiers
misc_feature	4379..5372
old_sequence	/*note="Specifically claimed region of intronless cDNA identified by exon trapping" replace(50652..50653, c)
old_sequence	/*tag="Changes Val codon to Leu codon" replace(50796..50797, c)
old_sequence	/*tag="c /*note="replaces Val codon by Leu codon" replace(5827..5828, c)
old_sequence	/*tag="d /*note="insertion, results in frameshift"
W09612033.1A1.	
15-AN-1995:	U13357
12-OCT-1994:	U5-323443.
31-JAN-1995:	U5-381520.
IG LAB INC.	100000000

Page 624

expression vector; baculovirus; ss.
 09-01-1987-A2.
 PPF 06-JUN-1985; IN5078.
 PPF 30-JUN-1985; GS-013420.
 PPR (NATU.) NATURAL ENVIRONMENT RES COUNCIL.
 PPR 04-JUL-1984; GS-013420.
 RFE 11-MAY-87; Lashop D., Possee R;
 RFE Wykes G.-0719809.
 DRD GENBANK; J22839.
 PT sequence - Californica nuclear polyhedrosis virus complete genome
 PT sequence - useful in the prodn. of vectors for enhanced
 PT heterologous protein expression, such as Interleukin(s),
 PS Interleukin(s).
 PS Disclosure; Page 90-186; 122pp; English.
 CCC The complete nucleotide sequence of the genome of clone 6 of the
 CCC baculovirus ATographa californica nuclear polyhedrosis virus (Acnpv)
 CCC J22839 was determined. The sequence is taken from the Genbank record
 CCC J22839.
 CCC Open reading frames (ORFs) 13, 20, 22-26, 28-33, 32-38, 41-46, 50-60,
 CCC 61-65, 67-70, 72-74, 76-78, 80-82, 84-86, 88-90, 92-94, 96-98, 100-102,

[illegible]


```
PT cds
PT      capsicum'
PT      /label= M2040
PT      /tag= P
PT      /note= 'Previously identified as MORP-19832 and
PT      MORP-20108, the encoded protein shows 41.92
PT      percent identity to spermidine/putrescine transport
PT      protein (potA) from E. coli'
PT      51525..52382
PT      /tag= Q2043
PT      /label= M2043
PT      /note= 'Previously identified as MORP-20110, the
PT      encoded protein shows 26.51 percentage
PT      identity to spermidine/putrescine transport
PT      protein (potB) from E. coli'
PT      52166..53220
PT      /tag= T
PT      /label= M2044
PT      /note= 'Previously identified as MORP-20111, the
PT      encoded protein shows 28.45 percentage
PT      identity to spermidine/putrescine transport
PT      protein (potC) from E. coli'
PT      54658..55605
PT      /tag= M2046
PT      /label= M2046
PT      /note= 'Previously identified as MORP-20112, the
PT      encoded protein shows 36.60 percentage
PT      identity to sialoglycoprotease (gcp)
PT      from Pasteurella hemolytica'
PT      58976..59310
PT      /tag= T
PT      /label= M2048
PT      /note= 'Previously identified as MORP-19834,
PT      MORP-20109, and MORP-20113, the encoded protein
PT      shows 43.02 percent identity to sialoglycoprotein
PT      recognition particle protein (fih) from B.
PT      subtilis'
PT      59117..59079
PT      /label= M2049
PT      /note= 'Previously identified as MORP-20114 and
PT      MORP-20115, the encoded protein shows 44.78
PT      percentage identity to purine nucleoside
PT      phosphorylase (deco) from E. coli'
PT      59083..59754
PT      /tag= V
PT      /label= M2050
PT      /note= 'Previously identified as MORP-20117, the
PT      encoded protein shows 81.03 percent identity to
PT      deoxyribose-phosphate aldolase
```

```
PT cds
PT      (deco) from Mycoplasma pneumoniae'
PT      complement (64898..65731)
PT      /tag= M2056
PT      /note= 'Previously identified as MORP-20122, the
PT      encoded protein shows 30.25 percent
PT      identity to the protein disclosed in
PT      G81D6185_99 from B. subtilis'
PT      65710..66249
PT      /tag= X
PT      /label= M2057
PT      /note= 'Previously identified as MORP-20123, the
PT      encoded protein shows 30.80 percentage
PT      identity to the protein disclosed in
PT      G81D6185_104 from B. subtilis'
PT      81047..82597
PT      /tag= Y
PT      /label= M2057
PT      /note= 'Previously identified as MORP-19845, the
PT      encoded protein shows 28.84 percentage
PT      identity to glutamic acid specific protease
PT      (gspA) from Staphylococcus aureus'
PT      91065..91919
PT      /tag= Z
PT      /label= M2070
PT      /note= 'Previously identified as MORP-20136, the
PT      encoded protein shows 30.80 percentage
PT      identity to ribonuclease protein S2 (rps2)
PT      from Spiroplasma plantensis'
PT      103104..104324
PT      /tag= aa
PT      /label= M2077
PT      /note= 'Previously identified as MORP-20140, the
PT      encoded protein shows 28.05 percentage
PT      identity to oligopeptide transport system
```

Note: remainder of annotations omitted.

Query Match 1.34; Score 15; DB 27; Length 580073;
Best Local Similarity 100.04; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 578014 atttattcccttgat 578028
Oy 642 CTTTATCCCTTGAAT 656

Search completed: Mon Nov 16 15:48:36 1998
Job time : 294 secs.

!!SEQUENCE LIST 1.0
(Nucleotide) WORDSEARCH of: /home/obryen/dra977/olig/us-08-887-977-9 check: 9278 from:
FROMIG of: /home/obryen/dra977/olig/US08887977.seq
sequence 9, application us/08887977
general information:
applicant: wang, wei
applicant: gish, kurt c.
applicant: schall, thomas j. . . .

TO: tags: Sequences: 2,336,638 Total-length: 917,356,764 November 16, 1998 17:59

cat & sts
Database Release Information:

GenBank_Tags, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998
GenBank, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998
EMBL_Tags, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998
EMBL, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998

Word-size: 15 Words: 39166 Diagonals: 2,560 Total-diagonals: 2,000,000,000
Integral-width: 1 Alphabet: 4 List-size: 50 CPU minutes: 315.75

Sequence	Strd	Diag	Score	Width	Documentation
GB_EST19:AI045155	-	85	23	1	AI045155 UI-R-CJ-kk-c-08-0-UI.s1 UI-R
GB_EST13:N83058	-	90	20	1	N83058 TgESTzy59g10.r1 TgrH Tachyzoit
GB_EST7:AA255136	-	199	19	1	AA255136 mz82f03.r1 Soares mouse NML
GB_EST7:AA254935	-	-203	19	1	AA254935 mz80h09.r1 Soares mouse NML
GB_EST7:AA245646	-	4	19	1	AA245646 mx01a03.r1 Soares mouse NML
GB_EST6:AA148128	-	-352	19	1	AA148128 z031a06.r1 Stratagene colon
GB_EST5:W84101	-	-682	19	1	W84101 T2925 MVAT4 bloodstream form c
GB_EST5:AA060911	-	-265	19	1	AA060911 mj88b01.r1 Soares mouse p3NM
GB_EST5:AA014588	-	85	19	1	AA014588 mg93a01.r1 Soares mouse embr
GB_GSS2:AQ003850	-	-1013	19	1	AQ003850 CIT-HSP-2288K15.TF CIT-HSP H
GB_EST19:AI047830	+	-152	19	1	AI047830 ud64b12.x1 Sugano mouse live
GB_STS:G39250	+	-35	19	1	G39250 Z20386 Zebrafish AB Danio rerio
GB_EST1:R36281	-	-873	18	1	R36281 y968c11.r1 Homo sapiens cDNA c
GB_EST4:ATTS3711	+	-533	18	1	Z35153 A. thaliana transcribed sequen
GB_EST2:R35058	-	-372	18	1	R95058 Yq44f03.r1 Homo sapiens cDNA c
GB_EST2:R83706	-	-354	18	1	R83706 Yq14e10.r1 Homo sapiens cDNA c
GB_EST13:AA531233	+	-430	18	1	AA531233 nj52e06.s1 NCI_CGAP.Pr9 Homc
GB_EST12:AA225058	-	-176	18	1	AA225058 nc21b12.r1 NCI_CGAP.Pr1 Homc
GB_EST5:W25156	-	-190	18	1	W25156 z569c07.r1 Soares fetal lung N
GB_GSS4:AQ227870	+	-203	18	1	AQ227870 HS_2020.B1.A02_MR CIT Approv
GB_GSS3:AQ224100	-	-136	18	1	AQ224100 HS_2011.A1.E01_MR CIT Approv
GB_GSS3:AQ193807	+	-619	18	1	AQ193807 CIT-HSP-2384O12.TF CIT-HSP H
GB_GSS3:AQ185889	+	-133	18	1	AQ185889 HS_2240.A2.C09_MF CIT Approv
GB_GSS3:AQ133524	-	-267	18	1	AQ133524 HS_3051.A1.B10_MR CIT Approv
GB_GSS3:AQ118433	+	-298	18	1	AQ118433 HS_3007.A2.F10_T7 CIT Approv
GB_GSS1:B33590	-	-129	18	1	B33590 CIT-HSP-2011i1.TR CIT-HSP Homc
GB_GSS1:B31064	-	-690	18	1	B31064 HS-1005-B1-D08-MR.abi CIT Huma
GB_EST20:AI005637	+	-498	18	1	AI005637 cv59b10.s1 Soares testis NHT
GB_EST18:AA919945	+	-789	18	1	AA919945 yy47g04.r1 Stratagene mouse
GB_EST1:T64598	+	102	17	1	T64598 yc25c12.s1 Homo sapiens cDNA c
GB_EST1:T57466	-	-584	17	1	T57466 Y556a12.s1 Homo sapiens cDNA c
GB_EST1:T02045	-	-121	17	1	T02045 W8ST02766 Caenorhabditis elega
GB_EST1:T01726	-	-182	17	1	T01726 W8ST02447 Caenorhabditis elega
GB_EST5:AA009280	+	-14	17	1	AA009280 mg96c01.r1 Soares mouse embr
GB_EST4:W99065	+	-174	17	1	W99065 mf89g02.r1 Soares mouse embryc
GB_EST4:W98929	+	-629	17	1	W98929 mf91c09.r1 Soares mouse embryc
GB_EST4:W54177	+	45	17	1	W54177 md13a05.r1 Soares mouse embryc
GB_EST4:W44267	+	-170	17	1	W44267 mc74g10.r1 Soares mouse embryc
GB_EST4:W03671	+	-113	17	1	W03671 mc65d01.r1 Soares fetal liver
GB_EST4:N79232	+	-551	17	1	N79232 zal2h09.r1 Homo sapiens cDNA c
GB_EST4:W03671	+	-113	17	1	W03671 mc65d01.r1 Soares fetal liver
GB_EST4:N79232	+	-551	17	1	N79232 zal2h09.r1 Homo sapiens cDNA c
GB_EST4:MMTEST675	-	-135	17	1	Z31244 M.musculus expressed sequence
GB_EST4:ATTS4275	+	-466	17	1	Z45531 A. thaliana transcribed sequen
GB_EST3:N55392	+	-587	17	1	N55392 yz18d11.r1 Homo sapiens cDNA c
GB_EST3:HSC10D122	-	62	17	1	Z41661 H. sapiens partial cDNA sequen
GB_EST2:R78075	+	-414	17	1	R78075 Y180e04.r1 Homo sapiens cDNA c
GB_EST2:H02568	+	-829	17	1	H02568 Y136d05.s1 Homo sapiens cDNA c
GB_EST2:H00287	+	-51	17	1	H00287 Y122h10.r1 Homo sapiens cDNA c
GB_EST1:T79571	-	-931	17	1	T79571 YG75h09.r1 Homo sapiens cDNA c
GB_EST13:N81275	-	54	17	1	N81275 TgESTzy25h12.s1 TgrH Tachyzoit
GB_EST13:N69138	-	253	17	1	N69138 TgESTzy33a11.r1 TgrH Tachyzoit

of sequences containing matches at least 15 nt long

ORIGIN

US-08-887-977-9-Inst

CONTACT: MARIA W/HEUSER B2F Project
WASHU-HEMI Homeless B2F Project
Washington University School of Medicine
744 S. Euclid Avenue
744 S. Euclid Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouster@wustl.edu
The information is being made available free through HEMI - contact the
IMAGE Consortium (image.llnl.gov) for further information.
MCI:265856

RESULT	10	AA254935	457 bp	mRNA
LOCUS				

EST 14-MAR-1997


```

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 938 0200
Fax: 301 938 0208
E-mail: mdadm@net.org
Internet: mdadm@net.org
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M3-21
Class: BAC
Location/Qualifiers
1. 506
/organism="Homo sapiens"
/contig="p10"
/vector="peloBAC11; Site_1_ HindIII; Site_2:
HindIII"
/AB_xref="taxon:9606"
/clone="7288X15"
/clone_lib="CIR-RSP"
/seq_type="Genomic"
/seq_id="1"
/seq_len="98 g 185 t
162 a 61 c
BASE COUNT
ORIGIN
Query Match 1.7%: Score 19, DB 27; Length 506;
Best Local Similarity 100.0%; Pred. No.250e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 CTCGATGATCACTGAGTAA 57
Cp 68 CTCGATGATCACTGAGTAA 50

RESULT 14
Locus AAL46128 576 bp mRNA EST 05-DEC-1996
COLLOR: r1 Strazenevskan (c937204) Homo sapiens cDNA clone
ACCESSION AF046128, mRNA sequence.
Locus Length 576 bp
RID G171327
KEYWORDS EST.
SOURCE
human;
organism
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (base 1 to 576)
Bouček, T., Ellington, K., Bouček, M.,
Hollender, J., Alex, N., Kuchta, T., M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, P., Trevaaskie, E.,
Waterson, J., Williamson, A., Wohlmann, P. and Wilson, R.
MASH-Merck EST Project
Unpublished (1995)
Contact: Wilson RK

```

```

Schellenberg,K., Stepien,M.Tan,P., Underwood,K., Moore,B.,
Theisinger,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterson R.
The Washo-HMMI Mouse EST Project
Unpublished (1996)

Contact: Maria M/Mouse EST project
Washo-HMMI Mouse EST Project
University of Medicine
4444 Regent Parkway, Box #501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousecenter@wustl.edu
For information regarding the availability of cDNA libraries from the
IMACS Consortium (info@imacs.llnl.gov) for further information.
MGI:923971
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality: custom primer used
Location/Qualifiers
1. .785
/organism='Mus musculus'
/vector='Osgan; liver; Vector: pMS18S-PJ3; Site 1: DraIII
(CACTGTC); Site 2: DraIII (CAGCATGTG); 1st strand cDNA
was primed with an oligo(pT) primer extended cDNA was
[ATCTGGTTTCTTTTTTAATT]; double-stranded cDNA was
ligated into pUC19 vector and sequenced by PCR amplified
and cloned into distinct DraIII sites of the pMS18S-PJ3
vector ('5' site CACTGTC, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed before ligation. The library was constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTCTCTCTTAAAGCTCG and 3' end
primer CAAGCCCACTGCATCAACA.
/cname='U0090'
/clones='1450655'
/clone_lib='Sugano mouse liver alla'
/sex='female'
/sex_age='adult'
/lab_pos='DRIUS'
191 a 212 c 185 g 196 t 1 others
BASE COUNT
ORIGIN
Query Match 1.79; Score 19; DB 18; Length 785;
Best Local Similarity 100.0%; Pred. No. 2.50e+06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 724 AACATCGTGCTGCCGTCTG 742
|||||

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 295

High quality sequence stops: 295

TITLE
The WashU-Merck EST Project
JOURNAL
Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project


```

EST.
KEYWORDS Homo sapiens
          Human.
ORGANISM Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
          Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 463).
AUTHORS Biller, B., Clark, N., Dubouquet, T., Ellington, K., Hawkins, M.,
          Holman, M., Hutman, M., Kucaba, T., Le-M, L., Lennon, C., Marra, M.,
          Parsons, J. R., Schaffinger, T., Tan, P., Trivelpiece, A.,
          Watson, G. R., Wallingford, M., Whittam, P., and Wilson, R.
TITLE Washo-Werk EST Project
COUNTRY Unpublished (1995)
CONTACT: Wilson RK
Washo-Werk EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box #501, St. Louis, MO 63108
TELEPHONE: 314-772-2000
FAX: 314-286-1800
E-mail: est@watson.wustl.edu
This clone is available royalty free through LBNL, contact the
IMAGe consortium (info@image.lbnl.gov) for further information.
Seq primer: mob RG6+PT
High quality sequence step: 374.
Location/Qualifiers
      /organism="Homo sapiens"
      /note="Organ; Lung; Vector: p773D (Pharmacia) with a
      modified polylinker; Site_1: Not I; Site_2: Eco RI; let
      strand cDNA was primed with a Not I - oligo(dT) primer.",
      /db_xref="GenBank:U00001"
      /topology=circular
      double-stranded cDNA was size selected, ligated to Eco RI
      adapters (Pharmacia), digested with Not I and cloned into
      the Not I and Eco RI sites of a modified p773 vector
      containing a lacZ gene. The resulting recombinant library
      normalization to a Cot = 5 library constructed by Bento
      Soares and M.Fatima Bonaldo. This library was constructed
      from the same retus as the fetal heart library, Soares
      et al., 1987.
      /library=BMNH19
      /accession=U00001
      /date="Dec 1995"
      /clone="308844"
      /clone_lib="Soares fetal lung BMHL19w"
      /dev_stage="19 weeks"
      /analysis="DH10S (ampicillin resistant)"
      163 a 41 -462 113 g 92 t
      BASE COUNT
      mrna
      ORIGIN
Query Match 1.6%; Score 18; DB 23; Length 462.
Best Local Similarity 100.0%; pred. No. 1.59e+04;

```

Tue Nov 17 08:55:30 1998

US-08-887-977-9. rat

Page 51

THE NOV 17 09:55:30 1998

112-09-997-977-9-754

CS 000000

```

PCR product subcloned into pamp10 by the UPG-cloning
method (Life Technologies). Average insert size 1500
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman.
/clone.lib="NCI_CGAP_Pri"
/clone="MAGE1008767"
/sex="Male"
/dev_stage="45 years old"
/sex_age="463 DH103
41 - 463 DH103
157 a 98 c 112 g 93 t 3 others

Query Match
Best Similarity 100.0%; Pred. No. 1.59e+04;
Best Local 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches

DD 422 AAAATCTGGCTGCAGGAT 439
|||||
CP 522 AAAATCTGGCTGCAGGAT 505

RESULT 26
LOCUS R28056 477 bp mRNA EST 25-APR-1995
DEFINITION YH56007.r1 Homo sapiens cDNA clone 134724 5'.
ACCESSION R28056
VERSION R28056.1
KEYWORDS EST, 303
FEATURES
human clone134724 library=Soares placenta HD2HP vector=pT73D
(Pharmacia) with a modified polylinker host=DH10B (ampicillin
resistant, E. coli strain "Fus1" (Fusaro))
observed primer=5' primer=3' Note 1 (site2260 nt female placenta
Not 1 oligo(dT) primer (5'..1st strand cDNA was primed with a
AAGTGGAGATATCGGCGGAGAGATATTTTTTTTTTTTTTTT 3'), double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
I, ligated into the pT73D vector (Pharmacia), the modified pT73
vector. Library went through one round of normalization. Library
constructed by Sento Soares and M.Felina Bonaldi.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
Primates; Hominidae; Homo.
1 (bases 1 to 477)
Hiller,L., Clark,N., Dubuque,T., Elliott,K., Hawkins,M.,
Holman,M., Rutnan,M., Kucaba,T., Le,M., Lennon,C., Marra,M.,
Muller,M., Rattiner,C., Rife,D., Soares,M., Soares,M., F.,
Trevaskis,E., Waterson,K., Williams,R.A., Woldman,P. and
Wilson,R.
The NANO-Merck EST Project
Unpublished (1995)

```

Contact:	Wilson MK
Accession:	NCI-6089
Project:	Washington University School of Medicine
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel:	314 286 1800
Fax:	314 286 1700
E-mail:	cstevenson.wustl.edu
High quality sequence stops:	233
Source:	IMAGE Consortium; LMLL
This clone is available royalty-free through LMLL; contact the	
IMAGE Consortium (genetics@lml.gov) for further information.	
FEATURES	
source	1. 477 /organism=Homo sapiens* /clone=J4741
BASE COUNT	109 A 116 G 155 T 12 others
ORIGIN	
Query Match	1.5%
Best Local Similarity	94.7%; Score 18; DB 14; Length 477;
Matches 18; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
DB	318 CONSUMCONCNCVWNTT 336
CP	551 CTTGGCGGCTGTATTTT 533
RESULT 27	
LOCUS	AAS31233 509 bp mRNA 20-AUG-1997
DEFINITION	n152e06.a1 NCI_CGAP_P9 Homo sapiens cDNA clone IMAGE:596130, mRNA sequence.
ACCESSION	U01512
KEYWORDS	EST
NID	423273939
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	1. (bases 1 to 509) GenBank accession number U01512 ncbi.nlm.nih.gov/ncitgap
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
Contact:	Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue procurement: W. Maxton Linahan, M.D., Rodrigo P. Chusqui, Ph.D., National Cancer Institute CDNA LIBRARY PREPARED BY: David B. Krizman, Ph.D.


```

info identifies a large number of male germ cell-specific sequence
tags
JOURNAL 9161. Reprint. 52 (1), 131-138 (1995)
MEDLINE 9527560
REFERENCES 1 (bases 1 to 287)
AUTHORS Boog, C
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1994) C.Boog, Laboratory of Molecular Genetics,
Karolinska Institute, 171 77 Stockholm, Sweden
FEATURES
SOURCE 1..287
/organism="Mus musculus"
/strain="CBA"
/submitter="Boog, C"
/dev="stage=Adult"
/tissue_type="Testis"
/clone_lib="7-EAP"
BASE COUNT 67 a 81 c 68 g 73 t
ORIGIN
Query Match 1..58: Score 17; DB 23; Length 287;
Nest Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 131 CTCCTCCATCTGCGCA 147
Gv 266 CTCCTCCATCTGCGCA 282

RESULT 36 702045 289 bp mRNA EST 10-NOV-1992
ACCESSION C92027466 Caenorhabditis elegans cDNA clone C92563 similar to
DEFINITION Troponin I, Conservative
Tropoin I
ACCESSION 727826
XREF 927826
KEYWORDS
SOURCE
NAME
ORGANISM
NAME
REFERENCE 1 (bases 1 to 289)
AUTHORS McCombie W.R., Keilly J.M., Aubin, L., Goscoches M.,
Vizecard M.G., Wu, A., Adams, M.D., Dubnick M., Kerlavage, A.R.,
Kerlavage, M., Fields, C.A.
JOURNAL Unpublished (1993)
CONTACT: Kerlavage AR
336 Institute for Genomic Research
401 Westwood Drive, Gaithersburg, MD 20878
Tel: 3018599036

```

```

source
1. 296
   /organism="Homo sapiens"
   /note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
   with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
   Site_3: Xba I; Site_4: Kpn I; Site_5: Sal I; Site_6: Pst I;
   Site_7: Bam HI; Site_8: Not I; Site_9: Xho I; Site_10:
   15' strand:GATCAGTAAATTACCTCA with a Pac I...cgtg(cg) primer
   3' strand:GATCAGTAAATTACCTCA with a Pac I...cgtg(cg) primer";
   /db_xref="taxon:9606";
   /clone="297409";
   /contig="11";
   /contig_map="Soares fetal liver spleen NWRFS";
   /dev_stage="20 week-post conception fetus";
   /lab_host="DH10B (ampicillin resistant)";
   cl. >296 89 a 67 c 71 g 68 t 1 others
BASE COUNT
ORIGIN
Query Match
100% Identical: 17; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
Matches: 17;
Db 27 TCGTCGTCTTGCCGC 43
QY 140 TCGTCGTCTTGCCGC 156
RESULT 38
DEFINITION
Ov44008.x1 301 bp mRNA EST 16-JUN-1998
Ov44008.x1 Soares testis_NWT Homo sapiens CDNA clone IMAGE:1539191
3' UTRNA sequence.
ACCESSION
AF017132
KEYWORDS
MIMID
NAMES
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryote; Chordata; Vertebrate; Mammalia; Eutheria;
Primate; Hominin; Hominoidea; Homo.
1 bases 1 to 301.
REFERENCE
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE
Tumor Gene Index
JOURNAL
Unpublished (1997)
CONTACT
Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Fax: (301) 496-1550
WWW
http://www.strausberg.nih.gov
CDNA
Library: National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
... ..

```



```

/organs=Homo sapiens".
polylinker_Site_1: Not I; Site_2: Eco RI; let strand cDNA
was prepared from human tonsillar cells enriched for
B-lymphocytes by Dr. Louis M. Staudt (NCI). David Allan
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer.
17. Double-stranded cDNA ligated to EcoRI sites
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
screened by Dr. Louis M. Staudt and M. Fatima Bonaldo.
/ddb_xref=taxon:9608
/c clones=DAGE:703960
/clone_lib=NCL-GCAP.Cal.*
/lab_pos=NHL Biomedical Center B cell*
/complement<1..>310)
      80 a 101 c 78 g 51 t
nra
BASE COUNT
ORIGIN
          15%: Score 17; Length 310;
Query Match
Best Local Similarity 100.0%; Pred. No. 8.0Se-O3:
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Ddb 79 CTTTCGCTGCAGCTGCG 95
|||||.....|
Oy 887 CTTTCGCTGCAGCTGCG 903

RESULT 42
LOCUS AA332497 317 bp mRNA EST 21-APR-1997
DEFINITION Embryo, 8 week I Homo sapiens CDNA 5' end, mRNA sequence.
ACCESSION U00001
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens.
Eukaryotes: Chordata:
Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea:
Homo.
Homo (bases 1 to 317)
Aman, J., Lee H.S., Kirchner, E.F., Weinstein K.G., Gockeay R.A.,
Bult, C.J., Lee H.S., Kirchner, E.F., Weinstein K.G., Gockeay J.D.,
White, O., Sutton, D., Blake, J.A., Brandon, R.C., Man-Wai, C.,
Clayton, R.A., Clifton, T., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Friedland, J.M., Fritch, A.L., Frischauf, A.E., Fulton, S.B.,
Glennard, J., Gupta, C.K., Hanna-Miller, C., Heffron, J.B.,
Kelley, J.E., Kelley, J.C., Liu, L., Marmoras, S.M., Merrick, J.M.,

```

1054515 5' similar to gb:M70104 Mouse MHC class I D-region cell
antigen (M0058);, mRNA sequence.

ACCESSION
NID
REMARKS
SOURCE
ORGANISM

EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrates; Mammalia; Eutheria;
Rodentia; Sciurognathi; Rodidae; Murinae; Mus.
1 (bases 1 to 327)

REFERENCE
AUTHORS
MARIA,M., HILLIER,L., ALLEN,M., BOYLES,M., DIETRICH,F., DUBOQUIE,I.,
SCHALL,S., KUCASZ,K., LAM,K., LE,M., MARTIN,J., MORRIS,H.,
SCHEIDT,R., SUTHERLAND,G., TAYLOR,T., THOMPSON,C., WATKINS,
THEISING,B., WYLLIE,T., LENNON,G., SOARES,B., WILSON,R. and
WATERSTON,R.

TITLE
GENERAL
COMMENT

The washU-HMI Mouse EST Project
Unpublished (1996)

Contact: Marra M/Mouse EST Project
Washington University EST Project
Washington University School of Medicine
744 Forest Parkway, Box 8501, St. Louis, MO 63108
Fax: 314 285 1800
Fax: 314 285 1810
Email: mousetest@wustl.edu

This clone is available royalty-free through LLNM ; contact the
IMAGC Consortium (info@imc.lln.gov) for further information.
Seq primer: -28ml 5' end from Amersham.

FEATURES
source

Location/Qualifiers
1..327
/organism='Mus musculus'
/note='Organ: mammary gland; Vector: pTV73-Pac
(Pharmacia) with a modified polyniker; Site.1: Not I;
Site.2: Eco RI; lat strand cDNA was primed with a Not I-
oligo(dT) primer 5'-GGGCGCCGCCGCCTGTCGTTGCTGTTTTTTTTTTT
T-3'; Double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTV73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bento.
/db_xref='taxon:10090'
/clone_lib='Soares mouse mammary gland KbMGMC'
/xref_mail='mammary gland'
/dev_stage='4 weeks.'
/lab_host='DH109'

[illegible]

```

BASE COUNT      61 a  98 c  104 g  64 t
ORIGIN
Query Match      1.5%; Score 17; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches          17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dn  33  ACCACAGGGGTCACCA  49
      |||.....|
Cp  917  ACCACAGGGGTCACCA  901

RESULT  44
LOCUS      AA951060      336 bp      mRNA      EST      18-MAY-1998
DEFINITION UI-R-21-fv-9-08-0-UI-R-21 UI-R-21 Rattus norvegicus CDNA clone
ACCESSION  AA951060
VERSION    1
KEYWORDS   EST  19932
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
REFERENCE  1. Karyura; Metaxos; Chordata; Vertebrata; Mammalia; Eutheria;
          1 (bases 1 to 336)
          2. Kari; Kurida; Murine; Mammalia; Eutheria.
          3. Bonaldo M.F., Lennon, G. and Soares M.B.
          4. Normalization and subtraction: two approaches to facilitate gene
          discovery
          5. Chomaz, Res. 6 (9), 791-806 (1996)
          6. 9704467

CONTACT: Soares, MB
Program for Rat Gene Discovery and Mapping
451 Hattori Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 9250
Fax: 319 335 9585
Email: moore@blue.wesg.uiowa.edu

The sequence tag present in the CDNA between the Morf site and the
conserved 3' UTR region of the R-21 CDNA is identical to the one from
the normalised adult 12-Day-Embryo library, CDNA library
Preparation: M. Fatma Bonaldo, Ph.D. Clone distribution: clones
will be available through Research Genetics
Seq primer: M13 Forward
              1. -336
              2. -336
              3. -336
              4. -336
              5. -336
              6. -336
              7. -336
              8. -336
              9. -336
              10. -336
              11. -336
              12. -336
              13. -336
              14. -336
              15. -336
              16. -336
              17. -336
              18. -336
              19. -336
              20. -336
              21. -336
              22. -336
              23. -336
              24. -336
              25. -336
              26. -336
              27. -336
              28. -336
              29. -336
              30. -336
              31. -336
              32. -336
              33. -336
              34. -336
              35. -336
              36. -336
              37. -336
              38. -336
              39. -336
              40. -336
              41. -336
              42. -336
              43. -336
              44. -336
              45. -336
              46. -336
              47. -336
              48. -336
              49. -336
              50. -336
              51. -336
              52. -336
              53. -336
              54. -336
              55. -336
              56. -336
              57. -336
              58. -336
              59. -336
              60. -336
              61. -336
              62. -336
              63. -336
              64. -336
              65. -336
              66. -336
              67. -336
              68. -336
              69. -336
              70. -336
              71. -336
              72. -336
              73. -336
              74. -336
              75. -336
              76. -336
              77. -336
              78. -336
              79. -336
              80. -336
              81. -336
              82. -336
              83. -336
              84. -336
              85. -336
              86. -336
              87. -336
              88. -336
              89. -336
              90. -336
              91. -336
              92. -336
              93. -336
              94. -336
              95. -336
              96. -336
              97. -336
              98. -336
              99. -336
              100. -336
              101. -336
              102. -336
              103. -336
              104. -336
              105. -336
              106. -336
              107. -336
              108. -336
              109. -336
              110. -336
              111. -336
              112. -336
              113. -336
              114. -336
              115. -336
              116. -336
              117. -336
              118. -336
              119. -336
              120. -336
              121. -336
              122. -336
              123. -336
              124. -336
              125. -336
              126. -336
              127. -336
              128. -336
              129. -336
              130. -336
              131. -336
              132. -336
              133. -336
              134. -336
              135. -336
              136. -336
              137. -336
              138. -336
              139. -336
              140. -336
              141. -336
              142. -336
              143. -336
              144. -336
              145. -336
              146. -336
              147. -336
              148. -336
              149. -336
              150. -336
              151. -336
              152. -336
              153. -336
              154. -336
              155. -336
              156. -336
              157. -336
              158. -336
              159. -336
              160. -336
              161. -336
              162. -336
              163. -336
              164. -336
              165. -336
              166. -336
              167. -336
              168. -336
              169. -336
              170. -336
              171. -336
              172. -336
              173. -336
              174. -336
              175. -336
              176. -336
              177. -336
              178. -336
              179. -336
              180. -336
              181. -336
              182. -336
              183. -336
              184. -336
              185. -336
              186. -336
              187. -336
              188. -336
              189. -336
              190. -336
              191. -336
              192. -336
              193. -336
              194. -336
              195. -336
              196. -336
              197. -336
              198. -336
              199. -336
              200. -336
              201. -336
              202. -336
              203. -336
              204. -336
              205. -336
              206. -336
              207. -336
              208. -336
              209. -336
              210. -336
              211. -336
              212. -336
              213. -336
              214. -336
              215. -336
              216. -336
              217. -336
              218. -336
              219. -336
              220. -336
              221. -336
              222. -336
              223. -336
              224. -336
              225. -336
              226. -336
              227. -336
              228. -336
              229. -336
              230. -336
              231. -336
              232. -336
              233. -336
              234. -336
              235. -336
              236. -336
              237. -336
              238. -336
              239. -336
              240. -336
              241. -336
              242. -336
              243. -336
              244. -336
              245. -336
              246. -336
              247. -336
              248. -336
              249. -336
              250. -336
              251. -336
              252. -336
              253. -336
              254. -336
              255. -336
              256. -336
              257. -336
              258. -336
              259. -336
              260. -336
              261. -336
              262. -336
              263. -336
              264. -336
              265. -336
              266. -336
              267. -336
              268. -336
              269. -336
              270. -336
              271. -336
              272. -336
              273. -336
              274. -336
              275. -336
              276. -336
              277. -336
              278. -336
              279. -336
              280. -336
              281. -336
              282. -336
              283. -336
              284. -336
              285. -336
              286. -336
              287. -336
              288. -336
              289. -336
              290. -336
              291. -336
              292. -336
              293. -336
              294. -336
              295. -336
              296. -336
              297. -336
              298. -336
              299. -336
              300. -336
              301. -336
              302. -336
              303. -336
              304. -336
              305. -336
              306. -336
              307. -336
              308. -336
              309. -336
              310. -336
              311. -336
              312. -336
              313. -336
              314. -336
              315. -336
              316. -336
              317. -336
              318. -336
              319. -336
              320. -336
              321. -336
              322. -336
              323. -336
              324. -336
              325. -336
              326. -336
              327. -336
              328. -336
              
```

110-00-007-077-0 not

```

FEATURES
  source
    1. 352
      /cylindrocapsa spiens*
      /cylindrocapsa spiens*
    55 a 97 c 104 g 92 t 4 others
    BASE COUNT
    ORIGIN
      Query Match 1.58; Score 17; DB 20; Length 352;
      Best Local Similarity 100.00; Pred. No. 8.05e-03;
      Hc101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      Db 60 TTCTTCGACTGCTGT 76
      Oy 889 TTCTTCGACTGCTGT 905
      RESULT 48

```

```

/lab-test=DH108"
100 a 41...253      90 g    82 t
ORIGIN
mrna
Base Count      100 a 41...253      90 g    82 t
BASE COUNT
Query Match      1.5% ; Score 17; DB 26; Length 333;
Best Local Similarity 100.0%; Pired.No. 8.05E-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      120 TGAAGATCTCGAAGC 136
          |||||||||
DY      956 TGAAGATCTCGAAGC 972

RESULT 49
LOCUS       H00287              353 bp     mRNA           EST
DEFINITION  Y12240.r1 Homo sapiens cDNA clone 149539 5'.
ACCESSION   U00287
KEYWORDS    EST.
SOURCE      human clone=19539 library=Soares placenta N62HP vector=PTT1D
            (Pharmacis) with a modified polylinker host=DH10B (ampicillin
            resistant) strain=JM109 (Kodak) plasmid=Kodac Plasmids
            obtained at birth (full term), lat strand cDNA was primed with a
            Not I - oligo(dT) primer [5].
            ACTGAGAATTGGCCGCCGCAGATATTTTTTTTTTTT 3'; double-stranded
            cDNA was ligated to EcoRI adapters (pharmacis), digested with Not
            I, ligated to BamHI adapters (pharmacis), digested with BamHI,
            sequenced with BamHI primers (pharmacis).
            vector. Library went through one round of normalisation. Library
            constructed by Sento Soares and M.Patima Bonaudo.
ORGANISM    Homo sapiens
REFERENCE   1. Soares, R., Lucena, J., Lucena, J., Lucena, J., Lucena, J.,
            Deuterostoma; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
            Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
            Euarchia; Archonta; Primates; Catarrhini; Homioidea; Homo.
            Miller, L., Clark, M., Dubouque, T., Elliott, K., Hawkins, M.,
            Holman, M., Fulton, M., Kubacka, T., Le, M., Lennon, C., Marra, M.,
            Parsons, J., Rifkin, L., Schilling, T., Soares, M., Tan, F.,
            Wierwakske, E., Waterson, R., Williamson, A., Wohlmann, P., and
            others.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
CONTACT    Contact: Wilson BK
COMMENT     WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box #501, St. Louis, MO 63108
            Tel.: 314 286 1800
            Fax: 314 286 1800
            Email: est@wustl.wustl.edu

```

```

/cdms-lib*/Toji Kohara unpublished cDNA library*
/dev stage="embryo"
/sex="hermaphrodite"
/tissue_type="embryo"
BASE COUNT      94 a      85 c      85 t
ORIGIN
Query Match
Query Ident Similarity 102.5a Score 17 DB 9 Length 360;
Pct Ident 100.0a Pct Ident 99.0a Pct Ident 99.0a
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 ACTGCTCTCGGAGAA 70
      |||||
      |||||
Tp 1056 ACTGCTCTCGGAGAA 1040
Research completed, Sat Nov 14 03:05:54 1998
Job time : 3240 mcs.

```

LOCUS	EST	RNA	EST
A0260146	353 bp	mRNA	19-MAR-1987
DESCRIPTION	va37e05.1 Soares mouse 3INR212 5 Mus musculus cDNA clone 733568 5'		
ACCESSION	mRNA sequence.		
NUMERIC	A0260146		
REPRODS	E2T		
ORGANISM	house mouse.		
SOURCE	Mus musculus; mitochondrionl. cytochromes; Metaxoa, Chordata; Vertebrata; Eutheria; Rodentia; Sciurogasthi, Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 353)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Schellenberg,K., Steptoe,M., Tan,P., Underwood,K., Moore,B., Treising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Watson,R.		
TITLE	Cloning of Mouse EST Project		
JOURNAL	(Unpublished 1996)		
COMMENT	Contact: Maria M/Mouse EST Project Washington University School of Medicine Department of Molecular Biology and Genetics 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel. 314 286 1800 Fax. 314 286 1810 This clone is available royalty-free through LBNL, contact the IMAGE Consortium (info@image.llnl.gov) for further information. MEI450016 Seq. Present: -28mb3 rev7 ST from Amersham High quality features present: 106. Location/Qualifiers 1..353 /organism='Mus musculus' /vector='pGEM-T' /prim3-Pac (Pharmacia) with a modified polynucleotide primer: polymer: Site_1: Site_2: Eco RI; let strand cDNA was primed with a Not I - oligo(GT) primer [5'-GTGGCTGTTTCATCAGGCGGTTGAGGCTGGGGAGGCTGAGG-3'] on total mouse RNA provided by Minotry, Wayne State Univ.; double-stranded cDNA was ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. The recombinant plasmid was sequenced and was constructed by Bento Soares and M. Patricia Bonaldo." /db_xref='taxon:10090' /_clone='733568' /_contig='Soares mouse 3INR212 5' /_seq_unknown' /_dev_stage='11.5dpc total fetus'		

US-08-887-977-9, zzt
 Nov 17 08:55:30 1998

```

High quality sequence stops. 250
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLML ; contact the
IMAGE Consortium (infoimage@llnl.gov) for further information.
FEATURES             LOCATION/Qualifiers
-----             -
SOURCE              1..353
                    /organism="Homo sapiens"
                    /clone="149539"
BASE COUNT          108 a   60 c   73 g   109 t   3 others
ORIGIN
...
Query Match          1.5%; Pred. No. DB 14; Length 353;
Best Local Similarity 100.0%; Pred. No. 8.05e+03;
Matches            17; Conservative 0; Mismatches 0; Gaps 0;
DB      299 ACATCCCTTTGGTCTT 315
        |||||
DY     248 ACATCCCTTTGGTCTT 264
...
RESULT 50
LOCUS       C50266                360 bp    mRNA               29-AUG-1997
DEFINITION C.elegans C clone yk479f5 : 5' end, single read, mRNA sequence.
ACCESSION   C50266
VERSION     1
KEYWORDS    EST, EST(expressed sequence tag),
to mRNA, clone-1b-Yuji Kohara unpublished embryo CDNA library
SOURCE      Eukaryotic egress
           Clutterhouse, Metasod, Nematoda; Secernentes; Rhaditida; Rhaditidae;
Rhadinophora; Rhaditoidae; Rhaditidae; Peloderinae; Ctenorhabditis.
ORGANISM    Kohara,T., Mochizuki,T., Tabara,H., Shin-I.T., Matenabe,H.,
           Ohshima,M., Sugita,A.H., Sugita,I.,
           Obara,W., Sugimoto,A., Iida,K. and Nishigaki,A.
REFERENCE   1 (sites)
           Kohara,T., Mochizuki,T., Tabara,H., Shin-I.T., Matenabe,H.,
           Ohshima,M., Sugita,A.H., Sugita,I., Obara,W., Sugimoto,A.,
           Iida,K. and Nishigaki,A. Expression map of the C.elegans genome
           Unpublished (1997)
TITLE       2 clones 1 to 360)
MISCINFO   Direct Submission
SUBMITTER  Submitted (23-JUL-1997) to the DDBJ/EMBL/GenBank databases. Yuji
Kohara, National Institute of Genetics, Gene Network Lab, Yata
111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@nig.ac.jp,
Tel:+81-41-5400000 Location/Qualifiers
source      1..360
            /organism="Caenorhabditis elegans"
            /strain="n2"
            /accession="F03930"
            /clone="yk479f5"
  
```

Page 4

```

99 6 1.6 3011 6 R31621 Hepatitis C virus (NC 1.80e+02)
100 6 1.6 5035 5 R23450 MH mutant porcine Fya 1.80e+02

ALIGNMENTS

RESULTS
AC AC W48086; standard; Protein: 365 AA.
DT DT 11-JUN-1998 (first entry)
DE Human dendritic cell chemokine receptor.
KW KW human dendritic cell chemokine receptor; MIP-1alpha; MIP-1beta;
OS OS receptor; dendritic cell; macrophage; inflammation; asthma.
PR PR Homo sapiens.
PR Key Location/Qualifiers
PR Misc_Difference /note="encoded by CAV"
PR W09801557-A3.
PR 15-JAN-1998. H10018
PR 04-JUN-1997; US-044593
PR 05-JUN-1996; US-675814.
PR 11-OCT-1996; US-028329.
PR (SCHE 4) SCHEDING COMP.
PR PA (SCHE 4) SCHEDING COMP.
PR W011981001-03. Vicari A, Wang W, Eletskii A;
DR W011981004/03.
DR N-P528: V15418.
DR Novel chemokines, e.g. thymus-expressed chemokine - used for
DR treating primary myeloid leukemia and inducing asthma.
DR Citing primary patent: 6,919,300 (US).
CC CC The present sequence represents human dendritic cell chemokine receptor.
CC CC Antibodies which bind to the protein can be used in detecting or
CC CC diagnosing various immunological conditions related to expression
CC CC of the protein.
CC CC Isolating DNA clones for the chemokines, especially from other
CC CC species. The chemokine can be used in the treatment of conditions
CC CC associated with abnormal physiology or development, including
CC CC inflammatory conditions such as asthma.
CC CC Sequence 365 AA.
Query Match
Best Local Similarity 100.0%; Score 365; DB 29; Length 365;
Residues 365; Conservative 0; Altimatches 0; Indels 0; Gaps 0;
db 1 mftptpkllcsllhtqltkrcycpcrcgssgpylyrlaylcvlgllgnlivi 60
Qy 1 MFTPTPKLLCSLLHTQLTKRCYCPCRCGSSGPTLYRLAYLCVLGLLGNLIVI 60
db 61 lfsfaytkarmcdwylmmadadlrlvlpfvtvhtgawfawfnatckllsglyafn 120

```

CC		the effects of chemotherapy.	
SC	Sequence	360 AA:	
	Query Match	2.7% Score 10; DB 19; Length 360;	
	Best Local Similarity 100.0%;	Pred. No. 2.09e-03;	
	Matches 10; Conservative	0; Mismatches 0; Indels 0; Caps 0;	
DB	71 rscstcwljn 80		
QY	69 RSMTOVLLLN 78		
RESULT	3		
ID	W02689	standard peptide; 312 AA.	
AC	C2-MOV-1996	(first entry) telokin-B receptor	
DT	12-MOV-1996	G-protein coupled receptor; ligand binding domain	
TT	G-protein coupled receptor; ligand binding domain		
KN	schizophrenia; dopamine; cAMP; adenosine; thrombin; adrenergic; opsin;		
KW	muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;		
KW	sensory cytomegalo virus; serotonergic.		
KW	adenosine; cAMP; adenosine; thrombin; adrenergic; opsin;		
KW	muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;		
KW	sensory cytomegalo virus; serotonergic.		
PR	DSS508384-A		
PR	DSS508384-A		
PD	16-APR-1996.		
PF	10-SEP-1992; 942336.		
PP	10-SEP-1992; 942336.		
PR	09-SEP-1993; US-118270.		
PA	(UNF) UNIV NEW YORK STATE.		
PI	Murphy RB, Schuster DI.		
DR	WPI; 96-208765/21	peptide - useful as antipsychotic agent, e.g.	
PT	for treating schizophrenia.		
PF	Disclosures: Column 103-106; 184pp: English.		
P6	Proteins W02657-W02730 represent a range of G-protein coupled receptor		
CC	GPRs) proteins selected from cAMP, adenosine, muscarinic acetylcholine,		
CC	opson, sensory cytomegalo virus, adenosine, thrombin, adrenergic, opsin;		
CC	odorant; cytomegalo viral and other GPR proteins. The receptors proteins		
CC	were used to design polypeptides, pref. based on the transmembrane		
CC	domains, for use in G-protein coupled receptor ligand binding assays.		
CC	GPRs having inhibiting effect in biological activity such as W02717-W02799		
CC	GPRs having stimulating effect in biological activity such as W02717-W02799		
CC	as examples of polypeptide fragments). The polypeptide fragments can		
CC	be used in compositions for treating subjects suffering from a pathology		
CC	related to a GPR abnormality e.g. a psychotic disorder such as		
CC	schizophrenia.		
SQ	Sequence 312 AA.		
Query Match	2.5% Score 9; DB 19; Length 312;		
Best Local Similarity 100.0%;	Pred. No. 4.46e-02;		
Matches 9; Conservative	0; Mismatches 0; Indels 0; Caps 0;		

DT	26-JUN-1993	(first entry)
DV	Human IL-8 receptor from clone p2.	
KW	Intracellular-8 receptor; probes; gene therapy; gro receptor;	
KW	intracellular-calcium mobilising; ligand-binding; MIP-2 receptor.	
DB	W93104739	
PD	W93104739	
PF	01-APR-1993	
PP	14-SEP-1992; 007641.	
PP	13-SEP-1991; 05-753548	
P1	LJSEF / MS DZPT HEALTH & HUMAN SERVICE.	
P1	Murphy PM	
DR	WPI: 93-117849/14.	
DR	N-PDB: Q38747.	
CC	The IL-8 receptor aminoacid sequence - and corresp. cDNA expressed in various cells or transfected cells,	
CC	for screening ligands of IL-8 receptor and gene therapy	
CC	Claim 1; Fig 3; 3pp; English.	
CC	cDNA libraries from 2 and 3.5 kb fractions of poly(A) ⁺ RNA from HEK6	
CC	libraries were screened with PCR oligonucleotide probe (from rabbit	
CC	IL-8 receptor) and under low stringency with a p2 cDNA probe	
CC	synthesised from random primers, to isolated the clone p2, encoding	
CC	coding IL-8R of related MIP-2 receptor may be detected using	
CC	portions of the p2 clone as probes. P2 may also be used to screen	
CC	for ligands of IL-8R and may also be used in gene therapy to treat a	
CC	patient deficient in IL-8R. The IL-8R is a gro receptor and has	
SQ	calcium-mobilising and ligand-binding properties. Sequence 355 AA.	
Query Match	2/54; Score 9; DB 7; Length 355; Exact Similarity 14; Indels 0; Gaps 0;	
Matches	9; Conservative 0; Mismatches 0; Gaps 0;	
DB	309 yafiqdfr 317	
QY	307 YAFIQDPR 315	
RESULTS		
AD	R65186 standard; Protein; 356 AA.	
AK	R65188;	
DT	19-APR-1995 (first entry)	
DE	Murine nu-subtype opioid receptor.	
DE	Nucleotide sequence; drug addition.	
FW	nu-subtype opioid receptor; MSOR; drug addition.	
FH	Key	
FF	Location/Qualifiers	
FT	modified_site 10..12	
FT	/note= 'Putative N-linked glycosylation site'	
FT	/xref= 'Threonine residue especially favourable	

Query Match 2.50: Score 9: DB 14: Length 400:


```

Oy 291 VTEVAFRCCLNPLVAFQCKPFRNFKLKLKLVKYSQPSGACRYSEISBQ 350
Db 355 TSTADNADNASSFTM 369
Oy 351 TSTADNADNASSFTM 365

RESULT 2
ENTRY 2
TITLE 2
ALTERNATE_NAMES 2
ORGANISM 2
DATE 2
ACCESSIONS 2
REFERENCE 2
AUTHORS 2
JOURNAL 2
TITLE 2
ACCESSION 2
KEYWORDS 2
SUMMARY 2
Query Match 2
Best Local Similarity 100.0%; Pred. No. 1.32e-08;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 87 LPVLTLPFAV 97
Oy 85 LPVLTLPFAV 95

RESULT 3
ENTRY 3
TITLE 3
ALTERNATE_NAMES 3
ORGANISM 3
DATE 3
ACCESSIONS 3
REFERENCE 3
AUTHORS 3
JOURNAL 3
TITLE 3
ACCESSION 3
KEYWORDS 3
SUMMARY 3
Query Match 3
Best Local Similarity 100.0%; Pred. No. 1.32e-08;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 87 LPVLTLPFAV 97
Oy 85 LPVLTLPFAV 95

RESULT 4
ENTRY 4
TITLE 4
ALTERNATE_NAMES 4
ORGANISM 4
DATE 4
ACCESSIONS 4
REFERENCE 4
AUTHORS 4
JOURNAL 4
TITLE 4
ACCESSION 4
KEYWORDS 4
SUMMARY 4
Query Match 4
Best Local Similarity 100.0%; Pred. No. 1.32e-08;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 87 LPVLTLPFAV 97
Oy 85 LPVLTLPFAV 95

```

```

#status preliminary: not compared with conceptual translation
#molecule_type mRNA
#residues 1-360 #label POW
#cross-references CB:020824; NID:695172; PID:695173
#note
GENETICS
#gene GDB:CKMR4
#cross-references CB:677463
#accession A53752
#molecule_type mRNA
#residues 1-360 #label POW
#cross-references CB:020824; NID:695172; PID:695173
#note
KEYWORDS
transmembrane protein
PEPTIDE
112-133
112-133
135-175
208-226
281-308
29-276,110-187
72,350
145
183,194
321
SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1.32e-06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 RSMVDVILN 80
Oy 69 RSMVDVILN 78

RESULT 4
ENTRY 4
TITLE 4
ALTERNATE_NAMES 4
ORGANISM 4
DATE 4
ACCESSIONS 4
REFERENCE 4
AUTHORS 4
JOURNAL 4
TITLE 4
ACCESSION 4
KEYWORDS 4
SUMMARY 4
Query Match 4
Best Local Similarity 100.0%; Pred. No. 1.32e-06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 RSMVDVILN 80
Oy 69 RSMVDVILN 78

```

```

#accession S55594
#status preliminary: nucleic acid sequence not shown
#molecule_type DNA
#residues 1-358 #label RTE
#cross-references CB:020824; NID:695172; PID:695173
#note
KEYWORDS
SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 86 GLGNLILVI 95
Oy 51 GLGNLILVI 60

RESULT 5
ENTRY 5
TITLE 5
ALTERNATE_NAMES 5
ORGANISM 5
DATE 5
ACCESSIONS 5
REFERENCE 5
AUTHORS 5
JOURNAL 5
TITLE 5
ACCESSION 5
KEYWORDS 5
SUMMARY 5
Query Match 5
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 310 YAFIQGKFR 318
Oy 307 YAFIQGKFR 315

RESULT 6
ENTRY 6
TITLE 6
ALTERNATE_NAMES 6
ORGANISM 6
DATE 6
ACCESSIONS 6
REFERENCE 6
AUTHORS 6
JOURNAL 6
TITLE 6
ACCESSION 6
KEYWORDS 6
SUMMARY 6
Query Match 6
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 310 YAFIQGKFR 318
Oy 307 YAFIQGKFR 315

```

```

#accession A53752
#status preliminary: not compared with conceptual translation
#molecule_type mRNA
#residues 1-358 #label RTE
#cross-references CB:020824; NID:695172; PID:695173
#note
KEYWORDS
SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 312 YAFIQGKFR 320
Oy 307 YAFIQGKFR 315

RESULT 7
ENTRY 7
TITLE 7
ALTERNATE_NAMES 7
ORGANISM 7
DATE 7
ACCESSIONS 7
REFERENCE 7
AUTHORS 7
JOURNAL 7
TITLE 7
ACCESSION 7
KEYWORDS 7
SUMMARY 7
Query Match 7
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 312 YAFIQGKFR 320
Oy 307 YAFIQGKFR 315

RESULT 8
ENTRY 8
TITLE 8
ALTERNATE_NAMES 8
ORGANISM 8
DATE 8
ACCESSIONS 8
REFERENCE 8
AUTHORS 8
JOURNAL 8
TITLE 8
ACCESSION 8
KEYWORDS 8
SUMMARY 8
Query Match 8
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 312 YAFIQGKFR 320
Oy 307 YAFIQGKFR 315

```

```

Tue Nov 17 08:55:21 1998

      211-234      $domain transmembrane $status predicted $label TM5
      251-271      $domain transmembrane $status predicted $label TM6
      308-328      $domain transmembrane $status predicted $label TM7
      $length 359  $molecular_weight 44425  $checksum 3617

Query Match
Best Local Similarity 100.0%: Pred. No. 3 90%-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 313 YAP1QKFR 321
QY 307 YAP1QKFR 315

      8
RESULT      type complete
ENTRY      $chemname (C-C) receptor 4 - mouse
TITLE      $chemical_name Mus musculus $common_name house mouse
ORGANISM   08-Mar-1996 sequence_revision 19-Apr-1996 $date_change
DATE      14-Dec-1997
ACCESSIONS $accession
REFERENCES  $authors Hoogerwerf, A.J.; Black, D.; Proudfoot, A.E.I.; Wells, T.N.C.;
            $journal Blochem Biophys. Res. Commun. [1996] 218:337-343
            $title      Binding of chemokines to murine and human CC chemokine
            $accession JG4587
            $molecular_type mRNA
            $length 1560 $label M00
            $residues $residue_number 1560; $ND:gl167851; PTD:el95632; PID:gl167852
            $experimental_source thymus

GENETICS
$gene
$enzyme
$feature
2,183,194
72,202,350
145
321
$length 300 $molecular_weight 41462 $checksum 7652
Query Match
Best Local Similarity 100.0%: Pred. No. 3 90%-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 72 SMDYVLN 80

```

```

Tue Nov 17 08:55:21 1998          US-08-897-977-10.rpx

      $molecule_type mRNA
      $residues 6-360 $label MUR
      $cross-references GB:M73969
COMMENT   This receptor, unlike IL8RA, binds several peptides besides
           interleukin-8, including GMO, MAV-2, and EMA-78.
FEATURES
     Region              GOB:IL8RB; IL8RA
     $cross-references GD:I27688; ONCM:146928
     $map_position 2435-2q35
KEYWORDS   G protein-coupled receptor; transmembrane protein
SOURCE     length 360 $molecular_weight 4759 $checksum 3062

Query Match                2.5% Score 9; DB 2; Length 360;
Match Local Similarity 100.0%; Pred. No. 3.90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db    314 YAFIQKFR 322      |||||
Qy    307 YAFIQKFR 315      |||||

RESULT 10      S65693      type complete
ENTRY         oploid receptor mu variant MOR1A - human
TITLE         Opoid Receptor Mu Variant MOR1A - Homo sapiens (common name)
AUTHOR        Li J, Wang X, Zhang L, Chen Z, Xu Q, Wu C, Zhou W,
DATE          12-Jul-1998 $sequence_revision 40-Jul-1998 $text_change
            10-Sep-1998
ACCESSIONS    S65693: 551216
REFERENCE     S65693
AUTHORS       L.A.; Mansoor, E.; Yang, D.
SUBMISSION    Related to the genome project, July 1994
DESCRIPTION   Expression of two variants of the human mu opiod receptor
$description   mRNA in SK-N-SH cells and human brain.

$accession    S65693
$definition    oploid receptor mu variant MOR1A
$molecule_type mRNA
$residue_type 1-392 $label BAR
$cross-references ENBL:U12569; NID:g607911; PID:g607912
REFERENCE     S5125
AUTHORS       Bare, L.A.; Mansoor, E.; Yang, D.
SUBMITTER     KBRF:U12569
SUBMISSION    Related to the genome project, July 1994
DESCRIPTION   Expression of two variants of the human mu opiod receptor
$molecule_type mRNA in SK-N-SH cells and human brain.
$accession    S5126
$definition    oploid receptor mu variant MOR1A
$molecule_type mRNA
$residue_type 1-392 $label RAW
$cross-references length 392 $molecular_weight 43919 $checksum 6577
SUMMARY

Query Match                2.5% Score 9; DB 2; Length 392;
Match Local Similarity 100.0%; Pred. No. 3.90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Db      332 CLNPVATP 340
Oy      301 CLNPVTAF 309

RESULT 11
ENTRY   mu-opioid receptor - rat
TITLE   26-Jul-1986 status Norwegian state Norway rat
        26-Jul-1986 sequence_refish 26-Jul-1986 seq_change
DATE    13-Mar-1997
AUTHORS J.P.; Zhang, G.; Bouvier, C.; Szec, C.; Konnekleiv,
        O.K.; Jerl, M.J.; Grandy, D.K.
ACCESSIONS A56517; I57951; A49680; I52334; S34593; A48799; I58154
REFERENCE 1. Neurochem. (1995) 64:14-24
JOURNAL Characterization and distribution of a cloned rat mu-opioid
        title
accession-references MID:I5096825
accession I56517
status preliminary; translated from GB/FBIU/DDBJ
molecule_type 1-338 label_RBS
accession-references EMBL:D02083; NID:G403573; PID:G403574
REFERENCE 197951
authors Chen, Y.; Nestek, A.; Vid, J.; Rurley, J.A.; Yu, L.
journal Mol. Cell. Neurosci. (1995) 8:14-24
title Molecular cloning and functional expression of a mu-opioid
        receptor from rat brain.
accession-references MID:I9341493
accession I57951
status preliminary; translated from GB/FBIU/DDBJ
molecule_type mRNA
accession-references 1-338 label_R2
accession-references GI:II3069; MID:G49250; PID:G49251
REFERENCES 1. Expier, C.M.; Holmes, J.D.; Wang, J.B.; Johnson, B.; Corbett,
        M.L.; Kuthan, D.R.; Ohn, G.R.; Linden, J.
        J. Biol. Chem. (1993) 268:26447-26451
JOURNAL Purification and partial amino acid sequence of a mu opioid
        title
accession-references MID:I407533
accession A49680
status preliminary
molecule_type cDNA
accession-references 272-291 label_EEP
**experimental_source brain membranes
**note
sequence extracted from NCBI backbone (NCBIP:I40841)
REFERENCE 15211
authors Kishi, M.; Roy, S.; Ramakrishnan, S.; Wride, R.; Loh, H.H.
journal Biochem. Biophys. Res. Commun. (1995) 209:561-574
```

Tue Nov 17 08:55:21 1998

[illegible]

Tue Nov 17 08:55:21 1998

```

R.H.
Proc. Natl. Acad. Sci. U.S.A. (1994) 91:9081-9085
Genomic structure and promoter sequence of a mouse mu
opioid receptor gene.
crossa-references EMBL:J557793
accession I48665
status translated from GB/EMBL/DBJ
molecule_type DNA
keywds glibal RES
crossa-references EMBL:J05561; XID:9555698; PID:956069
REFERENCE
149300
authors Rossi, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.
journal J Biol Chem
title Alternative splicing and a novel morphine-6
beta-glucuronide receptor.
crossa-references M0-1D:9537799
status 85634
molecule_type mRNA
keywds 1-398 glibal ROS
crossa-references EMBL:J05561; XID:9552310; PID:9505331
status 85634
status translated from EMBL Data
Library, November 1995
MOR-1 213/1; 386/3
keywds opiate receptors
alternative splicing; G protein-coupled receptor;
glycoprotein; transmembrane protein
length 398 smolecular-weight 44421 scheckum 8164
Query Match 2.59; Score 9; DB 2; Length 398;
Best Local Similarity 100.00; Post. No. 3 90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 330 CLUSTALW 308
QY 301 CLUSTALW 309
RESULT 14
ENTRY 156553
TITLE opiate receptor mu - human
ALTERNATE_NAMES mu-opioid receptor mu
AUTHOR KIMLIN
DATE 02-Jul-1996
SOURCE Genomaphila from Genomaphila
DATE 28-Aug-1997
ACCESSION 156553; X89991; 941075; 851215
REFERENCE 138593
AUTHOR A.; Burley, J.W.; Egel, L.S.; Chen,
Y.; Tian, M.; Liu, Y.; Spivey, H.;

```

#journal J. Neurosci. (1995) 15:2596-2606
#title The effect of chronic treatment with morphine on the regulation of functional
#description desensitization by calcium/calmodulin-dependent protein
#cross-references MOPD-95198115
#status kinase and protein kinase C.
#molecule_type mRNA
#residues 1-50, N', 52-233, V', 215-400 #label WAN
#cross-references GB:125119; PID:9452073
#accession M1075
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;
#submitters Griffon, C.A.; Uhl, G.R.
#submitters submitted to GenBank, August 1994
#accession A3899; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues 1-50, N', 52-233, V', 215-400 #label WAN
#cross-references GB:125119; PID:9452073
#accession M1075
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;
#submitters Griffon, C.A.; Uhl, G.R.
#submitters Human mu opiate receptor, cDNA and genomic clones,
#title FEBS Lett. (1994) 338:217-222
#description Pharmacologic characterization and chromosomal assignment.
#status S4107; nucleic acid sequence not shown
#molecule_type mRNA
#residues 5121-50, N', 52-400 #label WAN
#cross-references GB:137216; ONIM:600018
#accession B2115
#authors Barr, L.A.; Wansan, E.; Yang, D.
#submitters FEBS Lett. (1994) 354:215-218
#submitters Expression of two variants of the human mu opioid receptor
#title mRNA in SK-N-SH cells and human brain.
#accession S3113; preliminary
#molecule_type mRNA
#residues 387-400 #label BAR
#cross-references GB:OPRM1
#accession GDB:OPRM1
#map_position 624-6425
#keywords G protein-coupled receptor; glycoprotein; transmembrane
#protein
#feature
73-96 #domain transmembrane #status predicted #label TMD1
107-132 #domain transmembrane #status predicted #label TMD2
144-165 #domain transmembrane #status predicted #label TMD3
186-208 #domain transmembrane #status predicted #label TMD4
236-257 #domain transmembrane #status predicted #label TMD5

283-304 #domain transmembrane #status predicted #label TMD6
321-342 #domain transmembrane #status predicted #label TMD7
371-392 #domain transmembrane #status predicted #label TMD8
#cross-references MOPD-95198115
#status kinase and protein kinase C.
#molecule_type mRNA
#residues 1-50, N', 52-233, V', 215-400 #label WAN
#cross-references GB:125119; PID:9452073
#accession M1075
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;
#submitters Griffon, C.A.; Uhl, G.R.
#submitters submitted to GenBank, August 1994
#accession A3899; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues 1-50, N', 52-233, V', 215-400 #label WAN
#cross-references GB:125119; PID:9452073
#accession M1075
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;
#submitters Griffon, C.A.; Uhl, G.R.
#submitters Human mu opiate receptor, cDNA and genomic clones,
#title FEBS Lett. (1994) 338:217-222
#description Pharmacologic characterization and chromosomal assignment.
#status S4107; nucleic acid sequence not shown
#molecule_type mRNA
#residues 5121-50, N', 52-400 #label WAN
#cross-references GB:137216; ONIM:600018
#accession B2115
#authors Barr, L.A.; Wansan, E.; Yang, D.
#submitters FEBS Lett. (1994) 354:215-218
#submitters Expression of two variants of the human mu opioid receptor
#title mRNA in SK-N-SH cells and human brain.
#accession S3113; preliminary
#molecule_type mRNA
#residues 387-400 #label BAR
#cross-references GB:OPRM1
#accession GDB:OPRM1
#map_position 624-6425
#keywords G protein-coupled receptor; glycoprotein; transmembrane
#protein
#feature
73-96 #domain transmembrane #status predicted #label TMD1
107-132 #domain transmembrane #status predicted #label TMD2
144-165 #domain transmembrane #status predicted #label TMD3
186-208 #domain transmembrane #status predicted #label TMD4
236-257 #domain transmembrane #status predicted #label TMD5

Job time : 70 secs.

DR PROSITE: PS00137: G-PROTEIN RECEPTOR: 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 RN NON-TER 1 45
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSXEN 118 139 3 (POTENTIAL).
 FT TRANSXEN 140 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSXEN 161 180 4 (POTENTIAL).
 FT TRANSXEN 181 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSXEN 206 228 6 (POTENTIAL).
 FT TRANSXEN 229 270 5 (POTENTIAL).
 FT TRANSXEN 271 291 7 (POTENTIAL).
 FT TRANSXEN 292 312 2 (POTENTIAL).
 FT TRANSXEN 313 333 BY SIMILARITY.
 FT DISULFID 116 193 POTENTIAL.
 FT CARBOHYD 19 19
 FT NON-TER 353 353
 SQ SEQUENCE 353 AA: 39998 MW: 74564858 CRC32:
 Query Match 2.5%: Score 9; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 311 YAFICQKPR 319
 QY 307 YAFICQKPR 315
 RESULT 5
 ID IL8B_MOUSE STANDARD: PRT: 353 AA.
 AC Q28519;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CKCR-2) (FRAGMENT).
 GN IL8RB OR CKCR2.
 OS MACACA MULATTA (HESUS MACAQUE).
 OC EUTHERIA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN IL8RB_MOUSE STANDARD: PRT: 353 AA.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96175151.
 RA LOPEZ-LARREA C E., STIEN F., GONZALEZ S., GONZALEZ-ROYES S.,
 CC IMMUNOGNETICS 41:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 > IL-6 >
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -2- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -3- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: Z91116; E198176; .

RL J. BIOL. CHEM. 269:12391-12394(1994).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 > IL-6 >
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -2- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -3- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: L24445; G437652; .
 DR PIR: A53752; A53752.
 DR SCDB: GCR 08617; .
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 RN NON-TER 1 46
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSXEN 118 139 3 (POTENTIAL).
 FT TRANSXEN 140 160 4 (POTENTIAL).
 FT TRANSXEN 161 180 5 (POTENTIAL).
 FT TRANSXEN 181 205 6 (POTENTIAL).
 FT TRANSXEN 206 228 7 (POTENTIAL).
 FT TRANSXEN 229 270 8 (POTENTIAL).
 FT TRANSXEN 271 291 9 (POTENTIAL).
 FT TRANSXEN 292 312 10 (POTENTIAL).
 FT TRANSXEN 313 333 11 (POTENTIAL).
 FT DISULFID 117 194 POTENTIAL.
 FT CARBOHYD 19 19
 FT NON-TER 353 353
 SQ SEQUENCE 358 AA: 40632 MW: 9046808 CRC32:
 Query Match 2.5%: Score 9; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 312 YAFICQKPR 320
 QY 307 YAFICQKPR 315
 RESULT 7
 ID IL8B_MOUSE STANDARD: PRT: 359 AA.
 AC Q28519;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CKCR-2) (FRAGMENT).
 GN IL8RB OR CKCR2.
 OS MACACA MULATTA (HESUS MACAQUE).
 OC EUTHERIA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN IL8RB_MOUSE STANDARD: PRT: 359 AA.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96175151.
 RA LOPEZ-LARREA C E., STIEN F., GONZALEZ S., GONZALEZ-ROYES S.,
 CC IMMUNOGNETICS 41:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 > IL-6 >
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -2- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -3- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: L23637; G435094; .
 DR EMBL: L11239; G29366; .
 DR EMBL: D17630; G43672; .
 DR EMBL: L20337; G438801; .

DR PROSITE: PS00137: G-PROTEIN RECEPTOR: 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 RN NON-TER 1 45
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSXEN 118 139 3 (POTENTIAL).
 FT TRANSXEN 140 160 4 (POTENTIAL).
 FT TRANSXEN 161 180 5 (POTENTIAL).
 FT TRANSXEN 181 205 6 (POTENTIAL).
 FT TRANSXEN 206 228 7 (POTENTIAL).
 FT TRANSXEN 229 270 8 (POTENTIAL).
 FT TRANSXEN 271 291 9 (POTENTIAL).
 FT TRANSXEN 292 312 10 (POTENTIAL).
 FT TRANSXEN 313 333 11 (POTENTIAL).
 FT DISULFID 116 193 POTENTIAL.
 FT CARBOHYD 19 19
 FT NON-TER 353 353
 SQ SEQUENCE 353 AA: 39947 MW: 8592P64 CRC32:
 Query Match 2.5%: Score 9; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 311 YAFICQKPR 319
 QY 307 YAFICQKPR 315
 RESULT 6
 ID IL8B_MOUSE STANDARD: PRT: 358 AA.
 AC Q28519;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CKCR-2) (GRO/MGSA
 CC RECEPTOR).
 OS ORNITHOLAGUS CUNICULUS (RABBIT).
 CC EUTHERIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; LAGOMORPHA.
 RN IL8B_MOUSE STANDARD: PRT: 358 AA.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9417014.
 RA CHERNETT R.P., NELSON H., KOSLOSKY C.J., MORRISSEY P.J.,
 CC GENOMICS 18:410-413(1993).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 > IL-6 >
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -2- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -3- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: L23637; G435094; .
 DR EMBL: L11239; G29366; .
 DR EMBL: D17630; G43672; .
 DR EMBL: L20337; G438801; .

DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CKCR-2) (GRO/MGSA
 CC RECEPTOR).
 GN IL8RB OR CKCR2 OR CKCR2 OR CKCR16.
 OS MUS MUSCULUS (MOUSE).
 CC EUTHERIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA; ROSENTIA.
 RN IL8RB_MOUSE STANDARD: PRT: 358 AA.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9417014.
 RA CHERNETT R.P., NELSON H., KOSLOSKY C.J., MORRISSEY P.J.,
 CC GENOMICS 18:410-413(1993).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 > IL-6 >
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -2- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -3- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: L23637; G435094; .
 DR EMBL: L11239; G29366; .
 DR EMBL: D17630; G43672; .
 DR EMBL: L20337; G438801; .

DR PIR: A53677; A53677.
 DR GCRD9; GCR 0530; -
 DR MGD; MGI:105303; CMOAR2.
 DR MGD; MGI:99567; GPCRL6.
 DR PROSITE; PS00337; G-PROTEIN RECEPTOR. 1.
 DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 75 83
 FT TRANSMEM 84 104
 FT DOMAIN 105 119
 FT TRANSMEM 120 141
 FT DOMAIN 142 162
 FT TRANSMEM 163 182
 FT DOMAIN 183 207
 FT TRANSMEM 208 230
 FT DOMAIN 231 251
 FT TRANSMEM 252 273
 FT DOMAIN 274 296
 FT TRANSMEM 297 314
 FT DOMAIN 315 359
 FT DISULFID 118 196
 FT DISULFID 119 196
 SQ SEQUENCE 359 AA; 40532 MW; 135249A CRC32;
 Query Match 2.54; Score 9; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 313 YAFIQKQFR 321
 QY 307 YAFIQKQFR 315
 RESULT 8 STANDARD; PRT: 359 AA.
 AC P55407; 1994 (REL. 23, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MSEA
 DE RECEPTOR)
 OS HUMAN
 ON RATTUS NORVEGICUS (RAT)
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=LUNG.

RA GOBL A.E., WANG S., ZHOU Y., OEBERG K.;
 RA SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 RA (2)
 RA SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=LIVER;
 RA KOSHIMIZU K., SHIBATA P., WATANABE K., TSURUFUJI S., NAKAGAWA H.,
 RA FUJIOKA M.;
 RA SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RA (3)
 RA SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=SPLEEN, AND LUNG;
 RA DUNSTAN C.A.N., SALAFRANCA M.N., ADRIANI S., XIA Y., FENG L.,
 RA SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RL -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 G-PROTEIN THAT ACTIVATES PHOSPHATIDYLINOSITOL-DEPENDENT KINASE
 AND TO G-PROTEIN. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 AND TO GRO/MSEA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 G-PROTEIN THAT ACTIVATES PHOSPHATIDYLINOSITOL-DEPENDENT KINASE
 AND TO G-PROTEIN. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 AND TO GRO/MSEA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 DR EMBL; D61584; U95255;
 DR EMBL; D61584; U95255;
 DR PIR: A53677; A53677.
 DR GCRD9; GCR 0530; -
 DR MGD; MGI:105303; CMOAR2.
 DR MGD; MGI:99567; GPCRL6.
 DR PROSITE; PS00337; G-PROTEIN RECEPTOR. 1.
 DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 75 83
 FT TRANSMEM 84 104
 FT DOMAIN 105 119
 FT TRANSMEM 120 141
 FT DOMAIN 142 162
 FT TRANSMEM 163 182
 FT DOMAIN 183 207
 FT TRANSMEM 208 230
 FT DOMAIN 231 251
 FT TRANSMEM 252 273
 FT DOMAIN 274 296
 FT TRANSMEM 297 314
 FT DOMAIN 315 359
 FT DISULFID 118 196
 FT DISULFID 119 196
 SQ SEQUENCE 359 AA; 40532 MW; 3630134 CRC32;
 Query Match 2.54; Score 9; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 314 YAFIQKQFR 322
 QY 307 YAFIQKQFR 315

Query Match 2.54; Score 9; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 313 YAFIQKQFR 321
 QY 307 YAFIQKQFR 315
 RESULT 9 STANDARD; PRT: 360 AA.
 AC P55407; 1994 (REL. 23, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MSEA
 DE RECEPTOR)
 OS HOMO SAPIENS (HUMAN)
 ON HOMO SAPIENS (HUMAN)
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=LIVER;
 RA KOSHIMIZU K., SHIBATA P., WATANABE K., TSURUFUJI S., NAKAGAWA H.,
 RA FUJIOKA M.;
 RA SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RA (3)
 RA SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=SPLEEN, AND LUNG;
 RA DUNSTAN C.A.N., SALAFRANCA M.N., ADRIANI S., XIA Y., FENG L.,
 RA SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RL -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 G-PROTEIN THAT ACTIVATES PHOSPHATIDYLINOSITOL-DEPENDENT KINASE
 AND TO G-PROTEIN. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 AND TO GRO/MSEA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 G-PROTEIN THAT ACTIVATES PHOSPHATIDYLINOSITOL-DEPENDENT KINASE
 AND TO G-PROTEIN. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 AND TO GRO/MSEA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 DR EMBL; M71969; G1109591;
 DR EMBL; M71969; G1109591;
 DR EMBL; M71969; G1109591;
 DR EMBL; M71969; G1109591;
 DR EMBL; M71969; G1109591;
 DR PIR: A53677; A53677.
 DR GCRD9; GCR 0530; -
 DR MGD; MGI:105303; CMOAR2.
 DR MGD; MGI:99567; GPCRL6.
 DR PROSITE; PS00337; G-PROTEIN RECEPTOR. 1.
 DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 48
 FT TRANSMEM 49 75
 FT DOMAIN 76 84
 FT TRANSMEM 85 105
 FT DOMAIN 106 120
 FT TRANSMEM 121 142
 FT DOMAIN 143 163
 FT TRANSMEM 164 183
 FT DOMAIN 184 208
 FT TRANSMEM 209 231
 FT DOMAIN 232 251
 FT TRANSMEM 252 273
 FT DOMAIN 274 296
 FT TRANSMEM 297 316
 FT DISULFID 119 196
 FT DISULFID 119 196
 SQ SEQUENCE 360 AA; 40759 MW; 135249A CRC32;
 Query Match 2.54; Score 9; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 314 YAFIQKQFR 322
 QY 307 YAFIQKQFR 315

Query Match 2.54; Score 9; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 4.77e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 314 YAFIQKQFR 322
 QY 307 YAFIQKQFR 315
 RESULT 10 STANDARD; PRT: 360 AA.
 ID CKR4.MOUSE
 AC P51600;

CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-DEPENDENT KINASE VIA A
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
CC
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -2- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC ENBL: X91114; E198175.

DR PROSITE: PS00337: G-PROTEIN RECEPTOR: 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOKINE;
FT DOMAIN 1 1
FT TRANSNEM 41 45 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 46 72 1 (POTENTIAL).
FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 82 112 1 (POTENTIAL).
FT DOMAIN 113 117 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 161 180 4 (POTENTIAL).
FT TRANSNEM 181 225 4 (POTENTIAL).
FT TRANSNEM 226 228 1 (POTENTIAL).
FT DOMAIN 229 248 6 (POTENTIAL).
FT TRANSNEM 249 270 6 (POTENTIAL).
FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 292 312 1 (POTENTIAL).
FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 19 19 POTENTIAL.
FT NON_TER 353 353
SQ SEQUENCE 353 AA; 39919 MW; 4AP43113 CRC32;

Query Match 2.2%; Score 8; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.03e-03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 311 YAFIQQKF 318
|||
QY 307 YAFIQQKF 314

Search completed: Mon Nov 16 10:35:11 1998
Job time : 32 secs.

Query Match 3.04; Score 11; DB 13; Length 185;
Best Local Similarity 100.04; Pred. No. 1.30e-02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 88 LVLVLPNAY 98
85 LVLVLPNAY 95

RESULT 3 PRELIMINARY; PRT: 383 AA.
AC Q89609;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE G-PROTEIN COUPLED RECEPTOR;
DE EQUINE HERPESVIRUS TYPE 2 (EHV-2).
OC VIRIDAE: DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RP [1] SEQUENCE FROM N.A.
RC STRAIN=86/67;
EX MEDLINE: 95102501.
RA TELFORD E.A., WATSON M.S., AIRD H.C., PERRY J., DAVISON A.J.;
J. MOL. BIOL. 249:520-528(1995).
RP SEQUENCE FROM N.A.
RC STRAIN=86/67;
RA TELFORD E.A.R.; 1995; TO EMBL/GENBANK/DBJ DATA BANKS.
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U20824; G695173;
DR PFAM: PF00001; 7tm.1.
SQ SEQUENCE 383 AA; 43667 MW; 60F58F74 CRC32;

Query Match 2.78; Score 10; DB 14; Length 383;
Best Local Similarity 100.04; Pred. No. 2.24e-07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 86 GLLGNLTVI 95
51 GLLGNLTVI 60

RESULT 4 PRELIMINARY; PRT: 383 AA.
AC Q43234;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE MU-OPFID RECEPTOR.

OS CATOSTOMUS COMBOSORI (WHITE SUCKER).
OC EPIKARYOTA: METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RP SEQUENCE FROM N.A.
RC TISSUE=CNS;
RA WALSHON M.G., HARVEY R.J., GUTEN F.R., KRETHAMAP H.J., ZWIERS H.,
SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: Y10904; E1169510;
DR PFAM: PF00001; 7tm.1.
DE G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
SQ SEQUENCE 393 AA; 43332 MW; 00BDCD CRC32;

Query Match 2.58; Score 9; DB 13; Length 383;
Best Local Similarity 100.04; Pred. No. 6.48e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 326 CLNVLVAV 334
301 CLNVLVAV 309

RESULT 5 PRELIMINARY; PRT: 368 AA.
AC Q42444;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE INTERLEUKIN-6-LIKE RECEPTOR;
DE ONCORHINCHUS MYLISS (RAINBOW TROUT) (SALMO GAIRDNERI).
OC EPIKARYOTA: METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTICHTHYES; ACTINOPTERYGII; SALMOPIFORMES.
RP SEQUENCE FROM N.A.
RA ZOO J., DANIELS G.D., CUNNINGHAM C., SECORGES C.J.;
SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: A000359; E1193478;
DR PFAM: PF00001; 7tm.1.
SQ SEQUENCE 368 AA; 41523 MW; 56D3903B CRC32;

Query Match 2.28; Score 8; DB 13; Length 368;
Best Local Similarity 100.04; Pred. No. 1.30e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 134 MLLATCIS 141
124 MLLATCIS 131

RESULT 6 PRELIMINARY; PRT: 373 AA.
AC Q57585;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
OS BRACHYDONT RERIO (ZEBRAFISH) (DANIO RERIO).
OC EPIKARYOTA: METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RP [1] SEQUENCE FROM N.A.
RA BARRALLO A., GONZALEZ-SARMIENTO R., GARCIA-MALVAR F., RODRIGUEZ R.E.;
SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: A001596; E1217996;
DR PFAM: PF00001; 7tm.1.
DE G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
SQ SEQUENCE 373 AA; 42520 MW; E90D0477 CRC32;

Query Match 2.28; Score 8; DB 13; Length 373;
Best Local Similarity 102.04; Pred. No. 3.0e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 66 GLLGNLTV 73
51 GLLGNLTV 58

RESULT 7 PRELIMINARY; PRT: 377 AA.
AC Q3733;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE NPYVA
OS BRACHYDONT RERIO (ZEBRAFISH) (DANIO RERIO).
OC EPIKARYOTA: METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RP SEQUENCE FROM N.A.
EX MEDLINE: 98068842.
RA LINDSEY J., BERGLUND M.M., STARRACK P., SALAMECK E., GEHLERT D.R.,
J. CELL BIOL. 16:1357-1363(1997).
RP SEQUENCE FROM N.A.
RA RINGVALL M., BERGLUND M.M., LARHAMMAR D.;
BIOCHEM. BIOPHYS. RES. COMMUN. 241:749-755(1997).

DR EMBL: AF037400; G3098346;
SQ SEQUENCE 377 AA; 42901 MW; ABDD4A3A CRC32;

Query Match 2.28; Score 8; DB 13; Length 377;
Best Local Similarity 100.04; Pred. No. 1.30e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 55 LGLLGNLTV 62
50 LGLLGNLTV 57

RESULT 8 PRELIMINARY; PRT: 377 AA.
AC Q83207;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE G-PROTEIN COUPLED RECEPTOR;
DE MURINE CYTOMEGALOVIRUS.
OC VIRIDAE: DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
RP [1] SEQUENCE FROM N.A.
RC STRAIN=SMITH;
RA DAVIS-POINTER N.J., LITCH D.M., VALLY H., SHELLAN G.R., RAWLINSON W.D.,
BARRELL B.G., PARFELLS E.E.;
SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: L41868; G1048723;
DR PFAM: PF00001; 7tm.1.
SQ SEQUENCE 377 AA; 42205 MW; A0179AC CRC32;

Query Match 2.28; Score 8; DB 14; Length 377;
Best Local Similarity 100.04; Pred. No. 1.30e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 302 LNPVLVAV 309
302 LNPVLVAV 309

RESULT 9 PRELIMINARY; PRT: 3345 AA.
AC Q46074;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE COSMID 3088.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EPIKARYOTA: METAZOA; ARTHROPODA; INSECTA; DIPTERA.

RM SEQUENCE FROM N.A.
RA NURPHY L., HARRIS D., BARRELL B.;
RL SUBMITTED (DEC-1997) TO ENBL/GENBANK/DBJ DATA BANKS.
RN (2)
RS SEQUENCE FROM N.A.
RA BROSSE M.
RL SUBMITTED (FEB-1998) TO ENBL/GENBANK/DBJ DATA BANKS.
DR ENBL: AL009195; E1248585;
SQ SEQUENCE 3345 AA; 357761 MW; D1D576D2 CRC32;
Query Match 2.2%; Score 8; DB 5; Length 3345;
Best Local Similarity 100.0%; Pred. No. 1.30e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3390 NISQTSSE 2397
QY 346 NISQTSSE 353
RESULT 10 PRELIMINARY: PRT: 57 AA.
ID 016823
AC 016823;
DT 01-JAN-1998 (TREMBLEL. 05, CREATED)
DT 01-JAN-1998 (TREMBLEL. 05, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE ANDROPIN.
GN ANP.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUTAROTIA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RC STRAIN-B316;
RC SEQUENCE FROM N.A.
RX MEDLINE; 97476321.
RA CHAKRABARTY S.;
RA GENOTEC 147; 715; 574(1997).
DR ENBL: AF018988; G2582253;
SQ SEQUENCE 57 AA; 6185 MW; D3AC277 CRC32;
Query Match 1.9%; Score 7; DB 5; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.40e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 5 WUFLA 11
QY 250 WUFLA 256
RESULT 11 PRELIMINARY: PRT: 66 AA.
ID 069444
AC 069444;
DT 01-AUG-1998 (TREMBLEL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE PROTEIN EXPORT MEMBRANE PROTEIN SEC.
GN SECG.
OS AQUIFEX ABOLICUS.
OC EUBACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUNAY M., HUBER R.,
RA GENOTEC 192; 353-358(1998).
DR ENBL: A2000673; G2982840;
SQ SEQUENCE 100 AA; 10464 MW; 3E125DC CRC32;
Query Match 1.9%; Score 7; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 10 VIAVVL 16
QY 246 VIAVVL 252
RESULT 14 PRELIMINARY: PRT: 102 AA.
ID Q21833
AC Q21833; 1996 (TREMBLEL. 01, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE R075.12.
OS CANOCHABITIS ELEGANS.
OC EUTAROTIA; METAZOA; ACCELORATES; NEMATODA; SECERNENTIA; RHABDITIDA.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA AINSOUCO R.;

DT 01-AUG-1998 (TREMBLEL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE ORF1 PROTEIN
OS METHYLOCCUS CAPSULATUS.
OC PROKARYOTA; GRACILICUTES; SCOPOBACTERIA; AEROSIC RODS AND COCCI;
OC METHYLOCCACEAE.
RC STRAIN-BATH, ATCC33009;
RC SEQUENCE FROM N.A.
RA TIPPETT A., JAHKE L., PORALIA R.;
RA EPOCHM. 0999; ACTA 1391; 223-232(1998).
DR ENBL: AF009586; G2911803;
SQ SEQUENCE 66 AA; 7338 MW; BC14204F CRC32;
Query Match 1.9%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 9 LHTQLI 15
QY 15 LHTQLI 21
RESULT 12 PRELIMINARY: PRT: 69 AA.
ID 048874
AC 048874;
DT 01-JUN-1998 (TREMBLEL. 06, CREATED)
DT 01-JUN-1998 (TREMBLEL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLEL. 06, LAST ANNOTATION UPDATE)
DE PLASMA MEMBRANE H+ATPASE (FRAGMENT).
OS GOSSTPID HIRSIUTUM (UPLAND COTTON).
OC EUTAROTIA; PLANTA; EMBRIOPHYTA; ANGIOSPERMAE; DICOTYLEDONAE;
OC MALVALES; MALVACEAE.
RC STRAIN-CV, ACALA SJ-2;
RC SEQUENCE FROM N.A.
RA SHART L.B., VODDANI F., MAESHIMA M., WILKINS T.A.;
RA PLANT PHYSIOL. 0:0-0(1998).
DR ENBL: AF009586; G2911803;
SQ SEQUENCE 69 AA; 7098 MW; DC266510 CRC32;
Query Match 1.9%; Score 7; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 52 GLSVIIS 58
QY 52 GLSVIIS 58
RESULT 13 PRELIMINARY: PRT: 100 AA.
ID 066505
AC 066505;
DT 01-AUG-1998 (TREMBLEL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE PROTEIN EXPORT MEMBRANE PROTEIN SEC.
GN SECG.
OS AQUIFEX ABOLICUS.
OC EUBACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUNAY M., HUBER R.,
RA GENOTEC 192; 353-358(1998).
DR ENBL: A2000673; G2982840;
SQ SEQUENCE 100 AA; 10464 MW; 3E125DC CRC32;
Query Match 1.9%; Score 7; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 10 VIAVVL 16
QY 246 VIAVVL 252
RESULT 14 PRELIMINARY: PRT: 102 AA.
ID Q21833
AC Q21833; 1996 (TREMBLEL. 01, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE R075.12.
OS CANOCHABITIS ELEGANS.
OC EUTAROTIA; METAZOA; ACCELORATES; NEMATODA; SECERNENTIA; RHABDITIDA.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA AINSOUCO R.;

RM SEQUENCE FROM N.A.
RA WILSON R., AINSOUCO R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA JOHNSON J., COOPER J., COULSON A., CHAYTON M.,
RA DANKS J., HILLIER L., JIER M., JOHNSON L., JONES M., KERSHAN J.,
RA KIRSTEN J., LAISTER E., LATREILLE P., LIGHTING J., LLOYD C.,
RA MCMURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA SKORZENIAK J., SAUNDERS D., SHOWKIER R., SHALLOON N., SMITH A.,
RA VAUDIN M., VAUGHAN K., WATKINSON P., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOLDMAN P.;
RA NATURE 368; 32-38(1994).
DR ENBL: A2000673; G2982840;
SQ SEQUENCE 102 AA; 11354 MW; 49B2E7FC CRC32;
Query Match 1.9%; Score 7; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 IIAVVL 77
QY 247 IIAVVL 253
RESULT 15 PRELIMINARY: PRT: 108 AA.
ID 058483
AC 058483; 1998 (TREMBLEL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE 109AA LONG HYPOTHETICAL PROTEIN.
OS PROTEIN.
OC EUTAROTIA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOTAMA A., MAGAI Y.,
RA SAJAI M., OGURA K., OTEURA R., NAKAGAWA H., TAKAMURA M., OHFUE Y.,
RA FUNAHASHI T., TANAKA T., KUDO Y., YAMAGUCHI J., KUSUDA N., OGUCHI A.,
RA KATO K., OSHIRO T., NAKAMURA Y., MASUCHI T., SHIOTA H., KIKUCHI H.;
RA SUBMITTED (JUL-1997) TO ENBL/GENBANK/DBJ DATA BANKS.
DR ENBL: AB004931; D1027866;
SQ SEQUENCE 108 AA; 12115 MW; 959841BA CRC32;
Query Match 1.9%; Score 7; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RM SUBMITTED (AUG-1994) TO ENBL/GENBANK/DBJ DATA BANKS.
RN (2)
RS SEQUENCE FROM N.A.
RA WILSON R., AINSOUCO R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA JOHNSON J., COOPER J., COULSON A., CHAYTON M.,
RA DANKS J., HILLIER L., JIER M., JOHNSON L., JONES M., KERSHAN J.,
RA KIRSTEN J., LAISTER E., LATREILLE P., LIGHTING J., LLOYD C.,
RA MCMURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA SKORZENIAK J., SAUNDERS D., SHOWKIER R., SHALLOON N., SMITH A.,
RA VAUDIN M., VAUGHAN K., WATKINSON P., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOLDMAN P.;
RA NATURE 368; 32-38(1994).
DR ENBL: A2000673; G2982840;
SQ SEQUENCE 102 AA; 11354 MW; 49B2E7FC CRC32;
Query Match 1.9%; Score 7; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 IIAVVL 77
QY 247 IIAVVL 253
RESULT 15 PRELIMINARY: PRT: 108 AA.
ID 058483
AC 058483; 1998 (TREMBLEL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLEL. 07, LAST ANNOTATION UPDATE)
DE 109AA LONG HYPOTHETICAL PROTEIN.
OS PROTEIN.
OC EUTAROTIA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RC STRAIN-WF5;
RC SEQUENCE FROM N.A.
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOTAMA A., MAGAI Y.,
RA SAJAI M., OGURA K., OTEURA R., NAKAGAWA H., TAKAMURA M., OHFUE Y.,
RA FUNAHASHI T., TANAKA T., KUDO Y., YAMAGUCHI J., KUSUDA N., OGUCHI A.,
RA KATO K., OSHIRO T., NAKAMURA Y., MASUCHI T., SHIOTA H., KIKUCHI H.;
RA SUBMITTED (JUL-1997) TO ENBL/GENBANK/DBJ DATA BANKS.
DR ENBL: AB004931; D1027866;
SQ SEQUENCE 108 AA; 12115 MW; 959841BA CRC32;
Query Match 1.9%; Score 7; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.64e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;